AMERICAN JOURNAL OF PRIMATOLOGY
Official Journal of The American Society of Primatologists ASP

# Density and distribution of western chimpanzees around a bauxite deposit in the Boé Sector, Guinea-Bissau.

| | |
|---|---|
| Journal: | *American Journal of Primatology* |
| Manuscript ID | AJP-18-0127.R4 |
| Wiley - Manuscript type: | Research Article |
| Date Submitted by the Author: | n/a |
| Complete List of Authors: | Dias, Filipe; Universidade de Lisboa, Instituto Superior de Agronomia; Universidade do Porto, Centro de Investigação em Biodiversidade e Recursos Genéticos; Universidade de Lisboa, School of Agriculture<br>Wenceslau, José; Foundation Chimbo,<br>Miller, David; University of Saint Andrews, Centre for Research into Ecological & Environmental Modelling and School of Mathematics and Statistics<br>Marques, Tiago; University of Saint Andrews, Centre for Research into Ecological & Environmental Modelling and School of Mathematics and Statistics; Universidade de Lisboa Departamento de Estatistica e Investigacao Operacional, Departamento de Estatística e Investigação Operacional |
| Indicate which taxonomic group was the subject of your study (select all that apply or type another option):: | Apes (non-human), Pan troglodytes verus |
| Keywords: | western chimpanzee, Boé, bauxite mining, Guinea-Bissau, Density surface modelling |
| | |

**Title:** Density and distribution of western chimpanzees around a bauxite deposit in the Boé Sector, Guinea-Bissau.

**Running title:** Western chimpanzees of Boé Sector.

José F. C. Wenceslau[1] and Filipe S. Dias[2,3,4*], Tiago A. Marques[5,6] and David L. Miller[6]

**José F. C. Wenceslau and Filipe S. Dias should be considered joint first authors**

[1] Foundation Chimbo, Amsterdam, The Netherlands

[2] Centre for Applied Ecology “Prof. Baeta Neves” (CEABN – InBIO), School of Agriculture, University of Lisbon, Tapada da Ajuda, 1349-017 Lisbon, Portugal

[3] CIBIO/InBio, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

[4] CIBIO/InBio, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisbon, Portugal

[5] Centre for Research into Ecological & Environmental Modelling and School of Mathematics & Statistics, University of St Andrews, St Andrews, Scotland, United Kingdom

[6] Centro de Estatística e Aplicações da Universidade de Lisboa, Departamento de Biologia Animal, Faculdade de Ciências da Universidade de Lisboa, Portugal

*Correspondence to: Filipe Dias (fsdias@isa.ulisboa.pt), CIBIO - Instituto Superior de Agronomia Tapada da Ajuda), Tapada da Ajuda, 1349-017 Lisboa

**Research Highlights**

- Approximately 18 nest building western chimpanzees inhabit the surroundings of a bauxite deposit in the SW of Guinea-Bissau;

- The construction of a mine can have adverse direct and indirect effects on this population.

## Abstract

The Boé sector in southeast Guinea-Bissau harbors a population of western chimpanzees (*Pan troglodytes verus*) that inhabits a mosaic of forest and savanna. The Boé sector contains a substantial bauxite deposit in a region called Ronde Hill, and there are plans for the construction of a mine, which may endanger the chimpanzee population. In a one-week survey in May 2013, we used the standing crop nest counts method to obtain the number of chimpanzee nests and from that estimate the density and abundance of chimpanzees. We carried out five 1 km line transects that covered the bauxite deposit and surrounding valleys. We used density surface modeling to analyze habitat preferences, then predicted chimpanzee nest density and distribution based on environmental variables. We found the projected location of the mine partially coincides with an area of high predicted abundances of chimpanzee nests and is surrounded by highly suitable areas for chimpanzees (northeast and southwest). We conclude the mine could have significant direct and indirect effects on this population of chimpanzees whose impacts must be carefully considered and properly mitigated if the mine is built.


**Keywords:** western chimpanzee, Boé, bauxite mining, Guinea-Bissau, Density Surface Modelling

## 1. Introduction

Western chimpanzees (*Pan troglodytes verus*, Schwarz) are a subspecies of chimpanzee whose distribution ranges from tropical lowland forests in Liberia, Côte d'Ivoire, and Sierra Leone to savannas in Guinea, Guinea-Bissau, Senegal, and Mali, that can also inhabit some highly humanized agro-forestry systems in these regions (Kühl et al., 2017). Western chimpanzees are currently listed as Critically Endangered in the International Union for the Conservation of Nature's Red List (Humle et al., 2016). The population of western chimpanzees declined by 80% and lost 20% of its range from 1990 to 2014 (Kühl et al., 2017). The most significant losses occurred in Côte d'Ivoire, where the population declined by 90%, mostly due to deforestation, poaching, and infectious diseases (Campbell, Kuehl, N'Goran Kouamé & Boesch, 2008). In Senegal and Ghana, there are fewer than 1000 individuals (Kormos & Bakarr 2003; Danquah, Oppong, Akom & Sam, 2012) and in Benin, Togo and Burkina-Faso western chimpanzees are probably extinct (Ginn, Robison, Redmond & Nekaris, 2013; Khül et al., 2017).



In Guinea-Bissau, chimpanzees were declared extinct in 1988, but subsequent surveys found populations in the Quinara and Tombali regions (in the southwest) and in Medina do Boé (a sector south of the Gabu region; Gippoliti, Embalo & Sousa, 2003; Brugiere, Badjinca, Silva & Serra, 2009). No country-wide abundance estimates are available for Guinea-Bissau, but some surveys suggest the population may range between 600 and 1000 individuals (Gippoliti et al. 2003). A study in Lagoas de Cufada Natural Park, in Quinara region, estimated 137 individuals (95% CI: 51–390) (Carvalho, Marques & Vicente, 2013). In Southern

Cantanhez National Park, in the Tombali region, a study reported fewer than 100 chimpanzees (Sousa, Barata, Sousa, Casanova, & Vicente, 2011). In the Boé sector, Serra, Silva, & Lopes (2007) interviewed hunters and other knowledgeable locals and came to an estimate of 710 individuals. The main threat to western chimpanzees in Guinea-Bissau is habitat loss and fragmentation due to expanding plantations of banana, cashew, and other fruits (Gippoliti et al., 2003). Expansion of mining operations can also impact chimpanzees, as some studies conducted in other West African countries have suggested (Diallo, 2010; Humle et al., 2016). Mining operations can have direct and indirect impacts on great apes (Arcus Foundation, 2014). The construction of mines can cause habitat loss, and mining operations can cause water contamination and habitat degradation (Kusin et al., 2017, Mensah et al., 2015). The noise from mineral extraction can disturb apes and cause them to move to other areas, thus disrupting their behaviors and social structure. The construction of roads for transporting minerals and workers can cause habitat loss, fragmentation, and increase disturbance (Arcus Foundation, 2014, Carvalho et al., 2013, Gippoliti et al., 2003; Hockings & Humle, 2009). The influx of new workers brought to work on mines can increase bushmeat hunting (Laurence et al., 2005) and promote conversion of forest into agricultural areas to cultivate crops. Frequent contact between humans and chimpanzees can also increase the probability of transmission of diseases for which chimpanzees lack immunity, such as bacterial respiratory diseases (Köndgen et al., 2008) and Ebola (Arcus Foundation, 2014, Devos, Sanz, Morgan, Ononenga & Laporte, 2008).

The Boé sector is located in the southeast of Guinea-Bissau and presents the highest altitudes in the country. The region contains lateritic plateaus, mostly close to the border with Guinea, with considerable amounts of bauxite (Diallo, 2010). Ronde Hill is where bauxite prospecting first began in the 1970s by Russian investors. In 2008, Bauxite Angola S.A. continued prospecting in association with Compagnie Bauxite de Guinée and built a road in the region. This road connects the deposit with the Republic of Guinea and is meant to facilitate the transportation of machinery for bauxite exploitation (Wit, 2011). Mining has not started and is contingent on agreements between Bauxite Angola S.A. and the Guinea-Bissau government that include the improvement of transportation infrastructure. Mining would take place at the crest of the hill, an area important for maintaining water quantity and quality in the Jabere and Paramaka rivers and adjacent valleys (Wit, 2011). Since these valleys host a population of western chimpanzees (Wit, 2011), it is crucial to assess the distribution of chimpanzees to understand the possible effects of mining and to develop mitigation strategies.

Here we estimate the abundance and distribution of chimpanzee populations in Ronde Hill and adjacent valleys to assess the potential impacts of a bauxite mine. We 1) determined the density and abundance of nest building chimpanzees based on the distribution of nests and 2) analyzed the overlap between chimpanzee nests and the mining area to assess potential impacts.

## 2. Methods

### Study area

The survey was conducted over approximately 47 km$^2$, comprising Ronde Hill, which includes the prospected bauxite deposit, and the basins of the rivers Paramaka and Jabere rivers and its tributaries, Barquere, Gra, Jabeje, Mussa and Tuncotanca creeks (Fig. 1). This site is in the southern limit of the Boé sector, which is close to the border with the Republic of Guinea (11° 41' N, 13° 54' W). The nearest human settlements are the villages of Capebonde in Guinea-Bissau and Paramakadow and Paramakaley on the Guinean side of the border. Soils in Ronde Hill are shallow and mostly in the early stages of laterization.  As a consequence, savanna is predominant, and forests occur only where the topsoil layer is deeper than one meter and does not flood for prolonged periods (Wit & Reintjes, 1989).

### Ethics statement

The present study complies with the Principles for the Ethical Treatment of Non-Human Primates of the American Society of Primatologists. This research was also approved by Guinea-Bissau's *Instituto da Biodiversidade e das Areas Protegidas* (IBAP). Since the sampling methods we used did not require direct contact between researchers and chimpanzees, disturbance and health threats to chimpanzees were minimal.

**Estimating the abundance of chimpanzees**

Since directly counting chimpanzees is often impractical, surveyors usually use indirect methods. In our case this involved counting nests, which chimpanzees build using branches and leaves. Nests are relatively easy to detect, remain visible for weeks, months, or even years and can be counted with distance sampling techniques (Buckland et al. 2001, Thomas et al. 2010). Chimpanzee abundances can then be estimated by combining the density of nests with nest construction rates, nest decay rates, and the proportion of the population that builds the nests (see below).

We established five parallel transects (each 1 km, North-South orientation) that were spaced one kilometer apart and encompassed Ronde Hill and adjacent valleys. During the first week of May 2013, three people followed the Standing Crop Nest Count (SCNC) protocol (Spehar et al., 2010): they walked along each transect carrying a GPS device (Garmin eTrex 10) and recorded the coordinates of chimpanzee nests and the perpendicular distance between each nest and the transect with a measuring tape. The decay stage of each nest was recorded following the scale used by Plumptre & Reynolds (1997): 1- if the nest is still fresh and stable, with green leaves and feces or feeding signs underneath, 2- if it is still solid, but the leaves have signs of drying, 3- if the nest presents only dried leaves and/or is starting to lose its structure, and 4- if it lost every leaf but is still recognizable as a nest due to the presence of broken branches and twigs. The surrounding environment around each nest was also classified according to four categories: 1) "primary forest" for pristine forested habitats or forests in later successional stages, 2) "secondary forest" for agricultural land abandoned for longer than five years that present dense mid-story and is starting to regain canopy closure, 3) "fallow" for agricultural fields abandoned for less than four years or still active, and 4) "savanna" for open or sparsely arborized grasslands. Contrary to the work of Bryson-Morrison, Tzanopoulos, Matsuzawa & Humle (2017) in Bossou, Republic of Guinea, our classification of "primary forest" encompasses mature and riverine forests, our "secondary forest" category includes young secondary forests and our "fallow" class corresponds to all types of highly disturbed habitats they identified in their study.

Chimpanzees tend to build nests in groups (Ogawa, Idani, Moore, Pintea & Hernandez-Aguilar, 2007). As recommended by Buckland et al. (2001), we considered clusters of nests as our observation unit instead of individual nests. To create clusters, we grouped nests with the same age class that were within 20 meters of each other *post hoc*. Some studies have used thresholds of 50 meters (e.g., Morgan & Sanz, 2006; Sousa et al., 2011), but based on our observations in the field we decided to choose 20 meters to reduce the risk of grouping different clusters together (see Marchesi, Marchesi, Fruth & Boesch 1995, Ogawa et al. 2007, Kouakou, Boesch, & Kuehl 2009).

Since chimpanzees show marked preferences for nesting sites (Carvalho, Meyer, Vicente & Marques, 2015; Bryson-Morrison et al., 2017), we used Density Surface Modelling (DSM) to model the abundance of clusters of nests (Hedley & Buckland, 2004; Miller, Burt, Rexstad & Thomas, 2013) as a function of environmental covariates that include topographic variables, distance to rivers, roads and villages, percentage of cover of different land uses and Shannon-Wiener land-use diversity (Table 1). Each of the transects was split into five 200 meter segments for modelling. This is a two-stage approach that involves 1) fitting a detection function to the clusters of nests and using it to estimate abundances in transect segments with a Horvitz–Thompson-like estimator (Borchers, Buckland, Goedhart, Clarke, & Hedley, 1998) and 2) building a generalized additive model (Wood, 2017) to model estimated cluster abundances per transect segment as a function of environmental covariates.

We fitted uniform, half-normal and hazard-rate detection functions and included observation-level covariates that may have affected nest detection, such as nest cluster size, mean nest age class and land use cover (savanna, primary forest, secondary forest or fallows). In dense forests and areas with dense understory, nest detection can be lower. Observed distances were truncated at 50 meters based on the visual inspection of the detection function superimposed on a histogram of distances (Buckland et al., 2001) (Appendix 1). The goodness of fit of each detection function was assessed with the Cramer-von Mises test and the Kolmogorov-Smirnov test (Buckland et al., 2004). The best detection function was selected using the Akaike's Information Criteria (AIC). All calculations were performed in R 3.6 (R Core Team, 2019) using the package "Distance" version 0.9.8 (Miller, Rexstad, Thomas, Marshall & Laake, 2016).

We used Generalized Additive Models (GAMs) to model the abundance of clusters of nests. The expected abundance in each segment was modeled with Tweedie or negative binomial distribution as a function of several covariates.

GAMs were fitted with the R package "dsm" version 2.2.17 (Miller et al., 2013). Thin plate regression splines (Wood, 2003) were used as the basis for the model's smooth terms. The model is initiated by considering that the fit is extremely wiggly. Then the fitting procedure induces a penalization that essentially means the final wigglyness is driven by the data. (Wood, 2017). To minimize the effects of correlation among covariates, we considered only those variables with an individually significant association ($p<0.05$) with nest cluster abundance. Furthermore, we calculated variance inflation factors (VIF; Fox & Weisberg, 2010) and eliminated covariates with a VIF > 3. After fitting the model with all variables, we removed non-significant terms to reduce concurvity. Smoothness selection was performed via restricted maximum likelihood (REML). Smooth terms were selected using approximate p-values ($p<0.05$) and by adding an additional penalty that allowed each smooth term to be removed during model fitting (Marra and Wood, 2011). Spatial autocorrelation was assessed by examining a correlogram of deviance residuals. To validate the final models, we analyzed deviance residuals and checked for normal distribution and constant variance (Wood, 2017). To calculate the density of chimpanzees we divided the estimated nest density by the nest production rate and nest decay rate (Plumptre, 2003), following a formula modified after Kühl, Maisels, Ancrenaz & Williamson (2008):

$$\mathrm{D}_{weanedchimpanzees} = \frac{\mathrm{D}_{allnests}}{r \times \mathrm{t}}$$

Where $r$ is the estimated rate of nest production per individual per day and $t$ is the estimated mean life of a nest. Both values can be calculated only by performing detailed field studies and may vary between populations and geographic areas. Because of time constraints, we could not estimate these parameters in our study area, so we used estimates from other studies. For *r* we used 1.09 nests/day per individual from Plumptre & Reynolds (1997) in Budongo Forest Reserve, Uganda. For *t* we chose 194 days from Fleury-Brugiere & Brugiere (2010) in the Haut Niger National Park, Republic of Guinea. This estimate was considered the most suitable given the proximity to our study area and similarities in climate and vegetation. Unfortunately, these studies did not provide the variances for these parameters. Therefore the variances of chimpanzee densities will be underestimated.

To assess the potential impacts of the construction of the mine on chimpanzees, we used the density surface model to calculate the predicted abundance of nests in the study area. We combined uncertainty from the spatial model (GAM) with that of detectability (detection function) using the delta method (assuming independence between these two components) using "dsm.var.gam" from the R package "dsm" (Miller et al 2013). Finally, we analyzed the overlap between the bauxite deposit and the areas where the model predicts higher abundances of nests.

## 3. Results

We counted 608 nests during the surveys, which we grouped in 116 clusters. The number of nests per cluster averaged 5.2 ± SD 6.7.

**Detection function**

We selected a hazard-rate key function with cluster size as a covariate by AIC. The truncation distance for the detection function was 50 m and selected by comparing test statistics from the Cramer–von Mises and Kolmogorov–Smirnov goodness of fit tests. The average detection probability was 0.534, and the coefficient of variation was 0.068 (Fig. 2). A complete comparison of the detection functions can be found in the Supplementary Information (Table S1), along with all the R code required to reproduce our results. Figure 2 shows relatively few detections close to the transect, which was caused by lower detectability of nests in areas with dense forest or dense understorey. This did not have important effects on the fit of the detection function.

**Density surface models**

The density surface model with a Tweedie distribution provided the best fit for the data (see quantile-quantile plot, Fig. 3). The abundance of clusters of nests was higher in areas with a northwest exposure, closer to seasonal rivers, in areas with a low cover of savanna and with a high Shannon-Wiener diversity of land uses (Fig. 4).

**Estimated abundance of nests and chimpanzees**

The model predicted the occurrence of 3878 nests in the study area. The coefficient of variation from the GAM was 0.2481, and the coefficient of variation of the detection function 0.1271. The total coefficient of variation for the estimate was 0.2788 (calculated using the delta method). Following Equation 1, the estimated abundance of nest building chimpanzees in Ronde hill is $N$ = 18 (95% CI: 11-31).

This estimate corresponds to a density of 0.3898 individuals/km$^2$ (95% CI: 0.2280–0.6664).

**The overlap between chimpanzees' nests and the proposed mine**

Predicted abundances of nests are not very high (< 20 nests) at the top of Ronde hill, where the mine is going to be built (there is some overlap in the northwestern part) (Fig. 5). The overlap between areas with a high predicted abundance of nests (>40 nests/km$^2$) and the future area of the mine is 0.2 km$^2$.

## 4. Discussion

In this study, we estimated the distribution and abundance of chimpanzees with the standing crop nest counts method and compared it with the future location of a bauxite mine. Overall, the predicted abundances of nests in location of the mine were relatively low, which can probably be explained by the fact that the top of Ronde Hill is covered by savanna and devoid of suitable trees for building nests. Still, the northeastern part of the mine coincides with an area of high observed and predicted nest density (>40 nests/km$^2$), that also contains the only accessible year-round source of water in a 2 kilometer radius. This area is probably an essential refuge for western chimpanzees, which are already suffering from habitat loss due to agricultural pressure from the neighboring village of Capebonde.

We estimated the total abundance of nest building chimpanzees in the study areas was 18 (95% CI: 11-31), corresponding to 0.3898 individuals/km$^2$ (95% CI: 0.2280–0.6664). Camera traps active during fieldwork placed in the valley of the Jabere river during identified at least 18 weaned chimpanzees (JFCW et al.

unpublished data). Our estimate is within the range of estimates obtained in other studies that also used the standing crop nest counts method. In Senegal, Pruetz et al. (2002) estimated 0.13 individuals/km$^2$, in the Republic of Guinea Fleury-Brugiere & Brugiere (2010) estimated 0.87 individuals/km$^2$ (95% CI: 0.73 – 1.04) and in Lagoas de Cufada Natural Park in Guinea-Bissau Carvalho et al. (2013) found 0.22 individuals/km$^2$ (95% CI: 0.08 – 0.62).

The density surface model suggests that chimpanzees prefer to build nests in areas facing northeast, with higher Shannon-Wiener land use diversity, with low cover by savanna, and close to seasonal rivers. These results are in line with the findings from other studies, which suggest that western chimpanzees can tolerate some human disturbance (Brugiere et al., 2009; Bryson-Morrison et al., 2017) and inhabit mosaics containing savanna, riparian forests, dense forests and more open habitats (Carvalho et al., 2013, 2015). In Lagoas de Cufada Natural Park (Guinea-Bissau), Carvalho et al. (2015) found that chimpanzees prefer to build nests in dense forests, contrary to our findings. Dense forests in Ronde hill are often close to frequently used agricultural areas which are avoided by chimpanzees. This type of avoidance behavior has also been observed in the Republic of Guinea (Bryson-Morrison et al., 2017).

Because of logistical constraints, we could conduct only one survey. We suggest that future research in the study area should focus on analyzing how chimpanzees use habitats throughout the year. It would also be useful to determine whether the chimpanzees that occur in Ronde Hill are part of one or several communities, and whether these communities are connected to those in the

Republic of Guinea. This information would allow us to better understand and prevent the possible impacts of the construction of the mine on this population of western chimpanzees.

**5. Conclusion**

The results of the study show that only a small part of the proposed mine coincides with areas of high chimpanzee's nests abundance. This small area of overlap presents one of the highest abundances of nests in the whole study area (>40 nests/$km^2$). In the remaining area around the mine, predicted nest densities are low, which probably reflects the fact that it is currently covered by grassland savanna and does not contain trees suitable for building nests. The projected location of the mine borders two areas of high abundance of chimpanzee's nests (northeast and southwest), therefore it is likely to be used by chimpanzees. The data we gathered, combined with the existing knowledge on impacts of mining on great ape populations, suggests the construction of the mine is likely to have significant direct and indirect effects on this population of chimpanzees. We recommend that if the mine is approved, authorities should carefully consider direct and indirect impacts on this population of chimpanzees and implement appropriate mitigation and compensation measures.

## 6. Acknowledgments

We thank CHIMBO Foundation, especially the director Annemarie Goedmakers and the board advisor Piet Wit for providing logistical and financial support for this survey and for their suggestions that significantly improved this work. We further thank the five CHIMBO's village committee members of Capebonde (Amadou Camará, Mangabói Culubali, Mari Cante, Boibalo Bangura, and Ali Camará) for their help in data collection and colleagues Jitske Willemsen and Menno de Boer for their company and support during the field survey. We are also extremely thankful to Bauxite Angola S.A. for providing us with their prospecting data and for the meeting with JFCW to discuss possible outcomes of this research. This work was approved by the *Instituto da Biodiversidade e das Áreas Protegidas* (IBAP) of Guinea-Bissau and complied with the Principles for the Ethical Treatment of Non-Human Primates of the American Society of Primatologists (ASP). JFCW was supported by *Ciências sem Fronteiras* scholarship from *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), Brazil (grant number 221350/2012-8). FSD was funded by FEDER funds through the Operational Programme for Competitiveness Factors - COMPETE and by National Funds through FCT - Foundation for Science and Technology under the UID/BIA/50027/2013 and POCI-01-0145-FEDER-006821. TAM thanks partial support by CEAUL (funded by FCT - Fundação para a Ciência e a Tecnologia, Portugal, through the project UID/MAT/00006/2019). We also wish to thank the editor and the anonymous reviewers for their comments that significantly improved this manuscript.

## 7. References

Arcus Foundation. (2014). Extractive industries and ape conservation (state of the Apes). https://doi.org/10.1017/CBO9781107590274

Borchers, D. L., Buckland, S. T., Goedhart, P. W., Clarke, E. D., & Hedley, S. L. (1998). Horvitz-Thompson estimators for double-platform line transect surveys. Biometrics, 54(4), 1221–1237.

Brugiere, D., Badjinca, I., Silva, C., & Serra, A. (2009). Distribution of chimpanzees and interactions with humans in Guinea-Bissau and Western Guinea, West Africa. Folia Primatologica, 80, 353–358. https://doi.org/10.1159/000259335

Bryson-Morrison, N., Tzanopoulos, J., Matsuzawa, T., & Humle, T. (2017). Activity and habitat use of chimpanzees (*Pan troglodytes verus*) in the anthropogenic landscape of Bossou, Guinea, West Africa. *International Journal of Primatology*, *38*(2), 282–302. http://doi.org/10.1007/s10764-016-9947-4

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2001). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations.* New York: Oxford University Press.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2004). *Advanced Distance Sampling: Estimating Abundance of Biological Population.* New York: Oxford University Press.

Campbell, G., Kuehl, H., N'Goran Kouamé, P., & Boesch, C. (2008). Alarming decline of West African chimpanzees in Côte d'Ivoire. *Current Biology*, *18*(19), R903–R904. http://doi.org/https://doi.org/10.1016/j.cub.2008.08.015

Carvalho, J. S., Marques, T. A., & Vicente, L. (2013). Population status of *Pan troglodytes verus* in Lagoas de Cufada Natural Park, Guinea-Bissau. *PloS One*, *8*(8), e71527. http://doi.org/10.1371/journal.pone.0071527

Carvalho, J. S., Meyer, C. F. J., Vicente, L., & Marques, T. A. (2015). Where to nest? Ecological determinants of chimpanzee nest abundance and distribution at the habitat and tree species scale. *American Journal of Primatology*, *77*(2), 186–199. http://doi.org/10.1002/ajp.22321

Danquah, E., Oppong, S., Akom, E., & Sam, M. (2012). Preliminary survey of chimpanzees and threatened monkeys in the Bia-Goaso Forest Block in Southwestern Ghana. African Primates. *African Primates 7*(2), 163-174. https://doi.org/10.1017/S0030605313001191

Devos, L., Sanz, C., Morgan, D., Onononga, J., & Laporte, N. (2008). Comparing ape densities and habitats in Northern Congo: surveys of sympatric gorillas and chimpanzees in the Odzala and Ndoki regions, *American Journal of Primatology 13*, 1–13. http://doi.org/10.1002/ajp.20514

Diallo, M.S. (2010). Praticas das indústrias extractivas na Africa ocidental, Síntese comparativa de quatro estudos de casos (Guiné-Bissau, Guinée, Senegal, Serra Leoa). Gland, Switzerland: IUCN, Dakar, Senegal: ASAN, Bissau, Guinea-Bissau: AD, Conakry, Guinea: Guinée ecologie, Freetown, Sierra Leone: EFA.

ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Fleury-Brugiere, M.-C., & Brugiere, D. (2010). High population density of *Pan troglodytes verus* in the Haut Niger National Park, Republic of Guinea: Implications for local and regional conservation. *International Journal of Primatology*, *31*(3), 383–392. http://doi.org/10.1007/s10764-010-9391-9

Fox, J., & Weisberg, H. S. (2010). An R companion to applied regression (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Ginn, L., Robison, J., Redmond, I., & Nekaris, K. (2013). Strong evidence that the West African chimpanzee is extirpated from Burkina Faso. *Oryx, 47*(3), 325-326. doi:10.1017/S0030605313000434

Gippoliti, S., Embalo, D. S., Sousa, C. (2003). Guinea-Bissau. In: R. Kormos, C. Boesch, M. I. Bakarr, & T. M. Butynski (Eds.), West african chimpanzees. status survey and conservation action plan (pp. 55–61). Gland, Switzerland and Cambridge, UK: IUCN/SSC Primate Specialist Group (PSG).

Hedley, S. L., & Buckland, S. T. (2004). Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics*, *9*(2), 181. http://doi.org/10.1198/1085711043578

Hockings, K., & Humle, T. (2009). Best practice guidelines for the prevention and mitigation of conflict between humans and great apes. Gland, Switzerland: IUCN/SSC Primate Specialist Group (PSG).

Humle, T., Boesch, C., Campbell, G., Junker, J., Koops, K., Kuehl, H. & Sop, T. 2016. *Pan troglodytes ssp. verus* (errata version published in 2016). The IUCN Red List of Threatened Species 2016: e.T15935A102327574. http://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T15935A17989872.en.

Köndgen, S., Kühl, H., N'Goran, P. K., Walsh, P. D., Schenk, S., Ernst, N., … Leendertz, F. H. (2008). Pandemic human viruses cause decline of endangered great apes. Current Biology, 18(4), 260–264. https://doi.org/10.1016/j.cub.2008.01.012

Kormos, R., & Bakarr, M.I. (2003). Regional Summary. In: R. Kormos, C. Boesch, M. I. Bakarr, & T. M. Butynski (Eds.), West african chimpanzees. status survey and conservation action plan (pp. 27–29). Gland, Switzerland and Cambridge, UK: IUCN/SSC Primate Specialist Group (PSG).

Kouakou, C. Y., Boesch, C., & Kuehl, H. (2009). Estimating chimpanzee population size with nest counts: validating methods in Taï National Park. *American Journal of Primatology*, *71*(6), 447–57. http://doi.org/10.1002/ajp.20673

Kühl, H. S., Maisels, F., Ancrenaz, M., & Williamson, E. A. (2008). Best practice guidelines for surveys and monitoring of great ape populations. Gland, Switzerland: IUCN/SSC Primate Specialist Group (PSG).

Kühl, H. S., Kalan, A. K., Arandjelovic, M., Aubert, F., D'Auvergne, L., Goedmakers, A., … Boesch, C. (2016). Chimpanzee accumulative stone throwing. *Scientific Reports*, *6*, 22219. http://doi.org/10.1038/srep22219

481 Kühl, H. S., Sop, T., Williamson, E. A., Mundry, R., Brugière, D., Campbell, G., … Boesch, C. (2017). The Critically Endangered western chimpanzee declines by 80%. *American Journal of Primatology*, *79*(9). https://doi.org/10.1002/ajp.22681

485 Kusin, F. M., Rahman, M. S. A., Madzin, Z., Jusop, S., Mohamat-Yusuff, F., Ariffin, M., & Z, M. S. M. (2017). The occurrence and potential ecological risk assessment of bauxite mine-impacted water and sediments in Kuantan, Pahang,Malaysia. *Environmental Science and Pollution Research*, *24*(2), 1306–1321. http://doi.org/10.1007/s11356-016-7814-7

490 Laurance, W. F., Croes, B. M., Tchignoumba, L., Lahm, S. A., Alonso, A. Lee, M. E. & Ondzeano, C. (2006). Impacts of roads and hunting on central African rainforest mammals. *Conservation Biology, 20,* 1251–1261. https://doi.org/10.1111/j.1523-1739.2006.00420.x

494 Marchesi, P., Marchesi, N., Fruth, B., & Boesch, C. (1995). Census and distribution of chimpanzees in Côte d'Ivoire, *Primates*, *36*(*4*), 591 – 607. https://doi.org/10.1007/BF02382880

497 Marra, G., & Wood, S. N. (2011). Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis*, *55*(7), 2372–2387. http://doi.org/https://doi.org/10.1016/j.csda.2011.02.004

500 Mensah, A. K., Mahiri, I. O., Owusu, O., Mireku, O. D., Wireko, I., & Kissi, E. A. (2015). Environmental impacts of mining: a study of mining communities in Ghana. *Applied Ecology and Environmental Sciences*, *3*(3), 81–94. http://doi.org/10.12691/aees-3-3-3

Miller, D. L., Rexstad, E., Burt, L., Bravington, M. V., & Hedley, S. (2019). dsm: Density surface modelling of distance sampling data (Version 2.2.17). Retrieved from https://cran.r-project.org/web/packages/dsm/index.html

Miller, D. L., Burt, M. L., Rexstad, E. A., & Thomas, L. (2013). Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution*, *4*, 1001–1010. https://doi.org/10.1111/2041-210X.12105

Miller, D. L., Rexstad, E., Thomas, L., Marshall, L., & Laake, J. L. (2016). Distance sampling in R. *Journal of Statistical Software*. http://dx.doi.org/10.1101/063891.

Morgan, D., & Sanz, C. (2007). Best practice guidelines for reducing the impact of commercial logging on great apes in Western Equatorial Africa. Occasional paper of the IUCN Species Survival Commission, No 34, IUCN

Ogawa, H., Idani, G., Moore, J., Pintea, L., & Hernandez-Aguilar, A. (2007). Sleeping parties and nest distribution of chimpanzees in the savanna woodland, Ugalla, Tanzania. *International Journal of Primatology*, *28*(6), 1397–1412. http://doi.org/10.1007/s10764-007-9210-0

Plumptre, A. J. (2003). Censuing chimpanzees. In: A. J. Plumptre, D. Cox, & S. Mugume, (Eds.), The status of chimpanzees in Uganda. Albertine Rift Technical Report Series No. 2. Wildlife Conservation Society, New York.

R Core Team (2017). R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Serra, A., Silva, C., & Lopes, E. (2007). Étude de faisabilité du projet "Développement touristique de la Boé au profit de la conservation des Chimpanzés et des populations locales". Bissau, Guinea-Bissau: WWF, Chimbo Foundation, IUCN and IBAP.

Sousa, J., Barata, A. V, Sousa, C., Casanova, C. C. N., & Vicente, L. (2011). Chimpanzee oil-palm use in southern Cantanhez National Park, Guinea-Bissau. *American Journal of Primatology*, *73*(5), 485–97. http://doi.org/10.1002/ajp.20926

Spehar, S. N., Mathewson, P. D., Wich, S. A., Marshall, A. J., Kühl, H., & Meijaard, E. (2010). Estimating orangutan densities using the standing srop and marked nest count methods: lessons learned for conservation. *Biotropica*, *42*(6), 748–757. https://doi.org/10.1111/j.1744-7429.2010.00651.x

Wit, P. (2011). Démarrage des activités d'exploitation de bauxite dans la Boé par l'entreprise "Bauxite Angola", Daridibó-Chimbo, Bissau.

Wit, P., & Reintjes, H. C. (1989). An agro-ecological survey of the Boé province, Guinea Bissau. *Agriculture, Ecosystems & Environment*, *27*(1), 609–620. http://doi.org/https://doi.org/10.1016/0167-8809(89)90121-7

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B*, *65*(1), 95-114. https://doi.org/10.1111/1467-9868.00374

Wood, S. N. (2017). Generalized Additive Models: An Introduction with R (2nd ed.). Boca Raton, FL: Chapman and Hall/CRC.

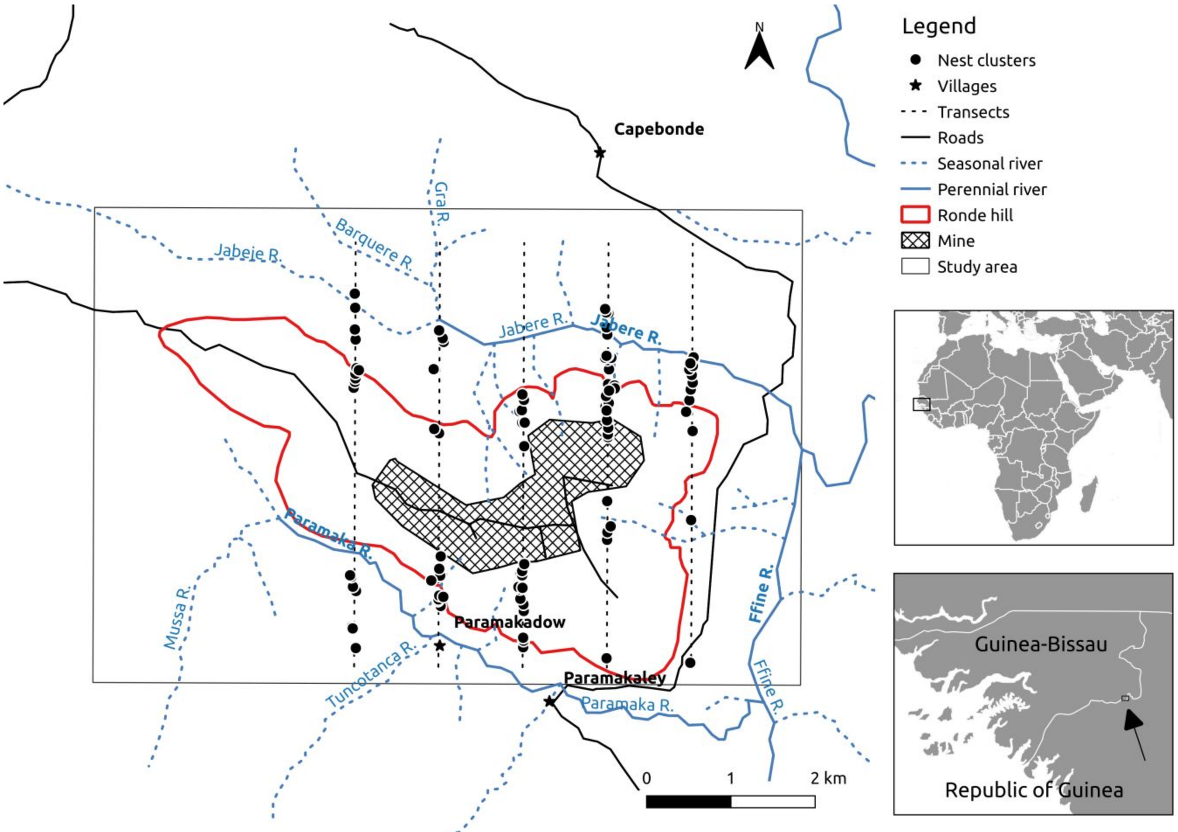





Figure 1 - Map showing the study area including the location of Ronde hill, the future location of the mine, transects, nest clusters, roads, rivers and closest villages. The top inset shows the location of Guinea-Bissau and the bottom inset the location of the study area in this country.

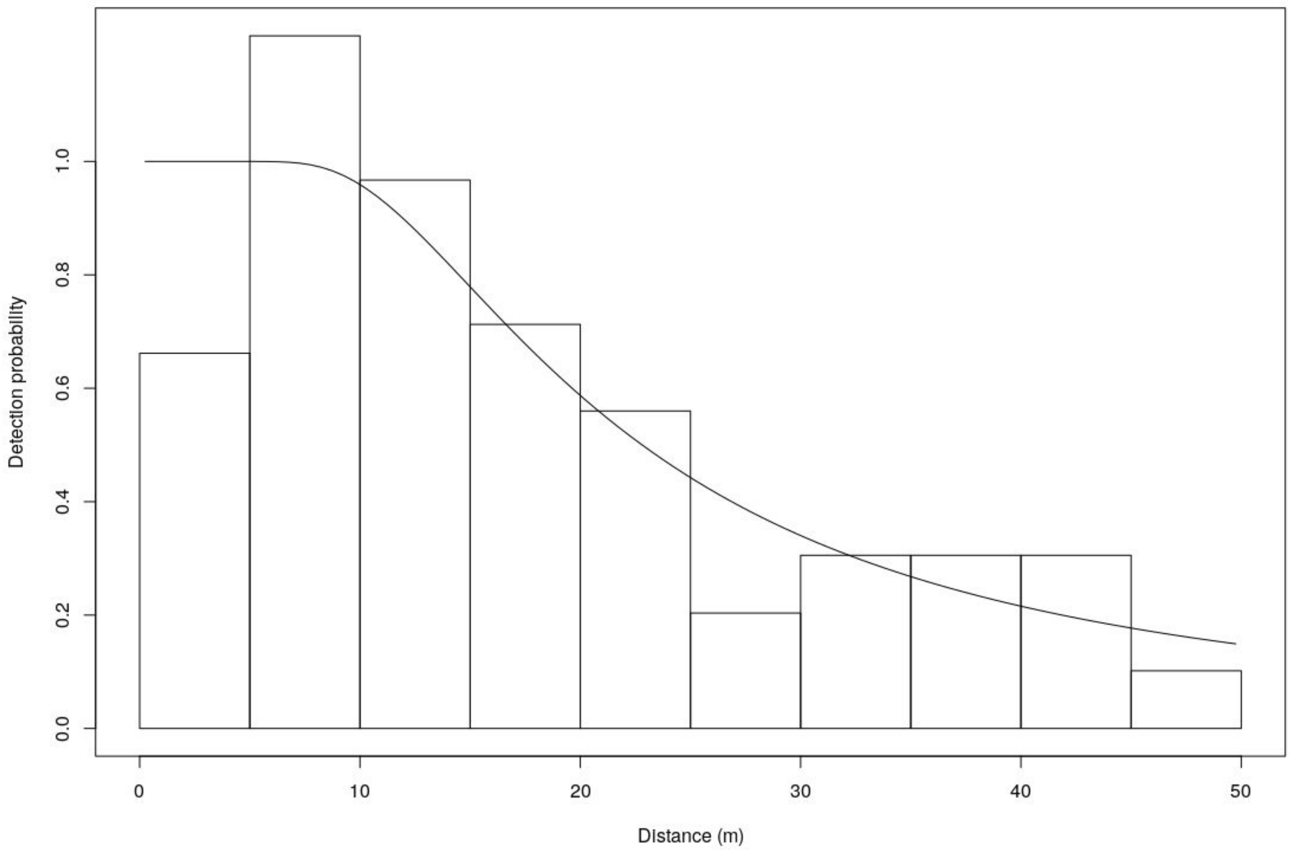



Figure 2 - Selected detection function (hazard-rate with cluster size as covariate) for clusters of nests overlaid onto a histogram of observed distances.

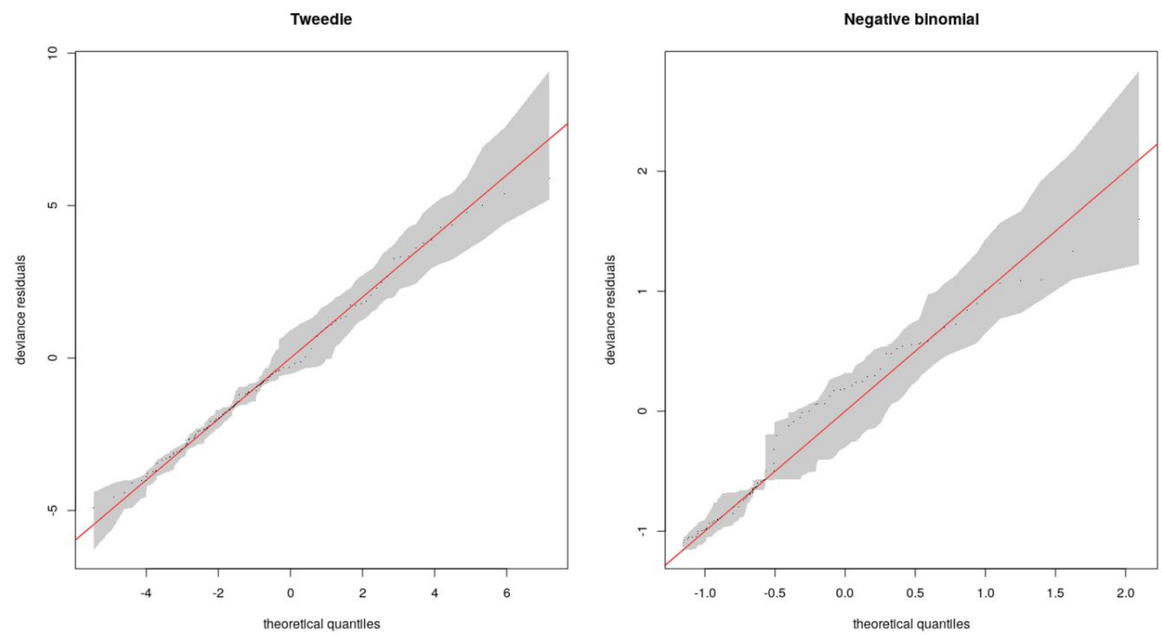



Figure 3 - Comparison of models with Tweedie (left) and negative binomial (right) response distributions by quantile-quantile plots. Good fit is indicated by agreement between observed and fitted (residual) quantiles (i.e., points being close to the red line). 90% reference bands are shown in grey allowing judgment of the deviation from the line. The negative binomial points fall further away from the red line than those for the Tweedie, indicating model misspecification.

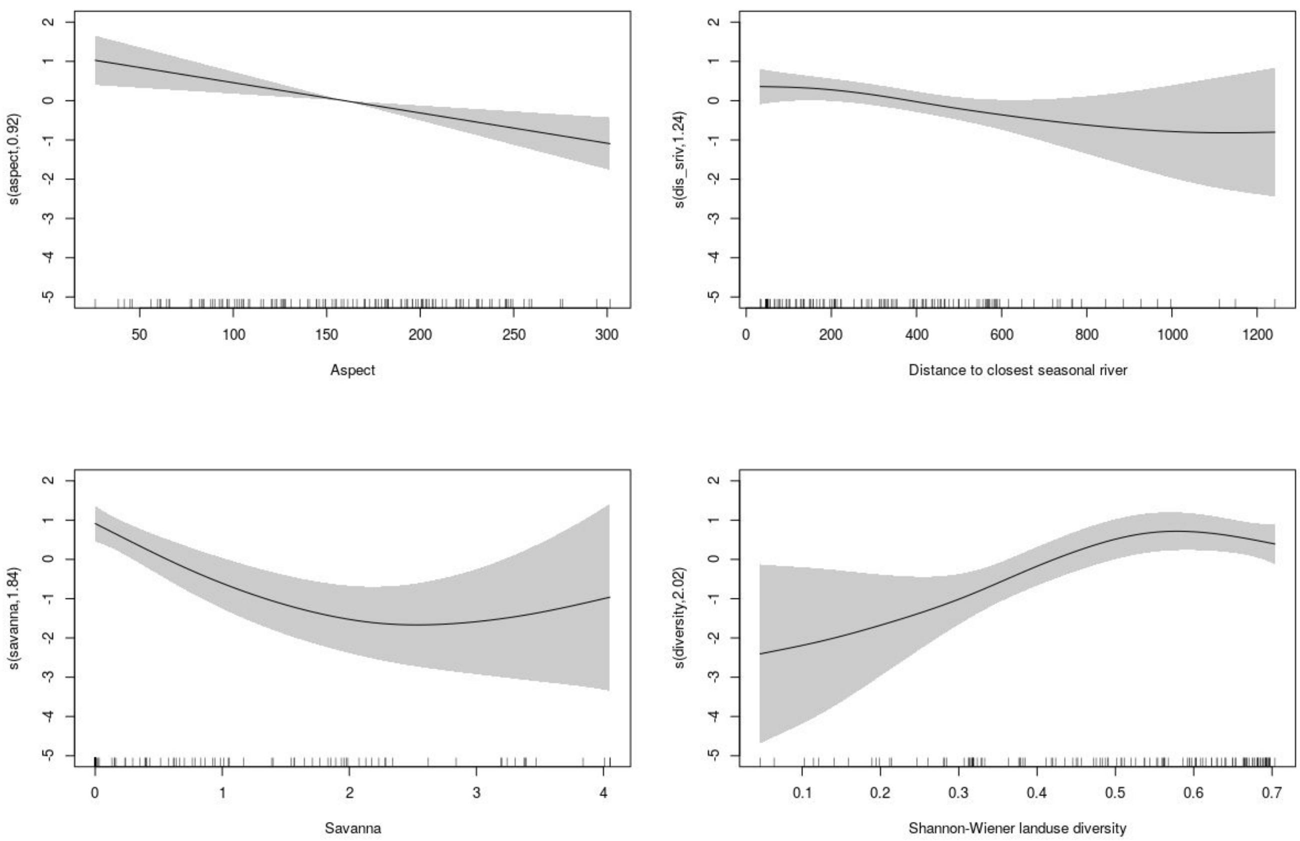



Figure 4 – Smooth functions for “aspect”, “distance to closest seasonal river”, “savanna and “Shannon-Wiener” land use diversity. Grey shading corresponds to 95% confidence bands, numbers in brackets on the vertical axis labels give the effective degrees of freedom of the term (1 corresponds to a linear term).

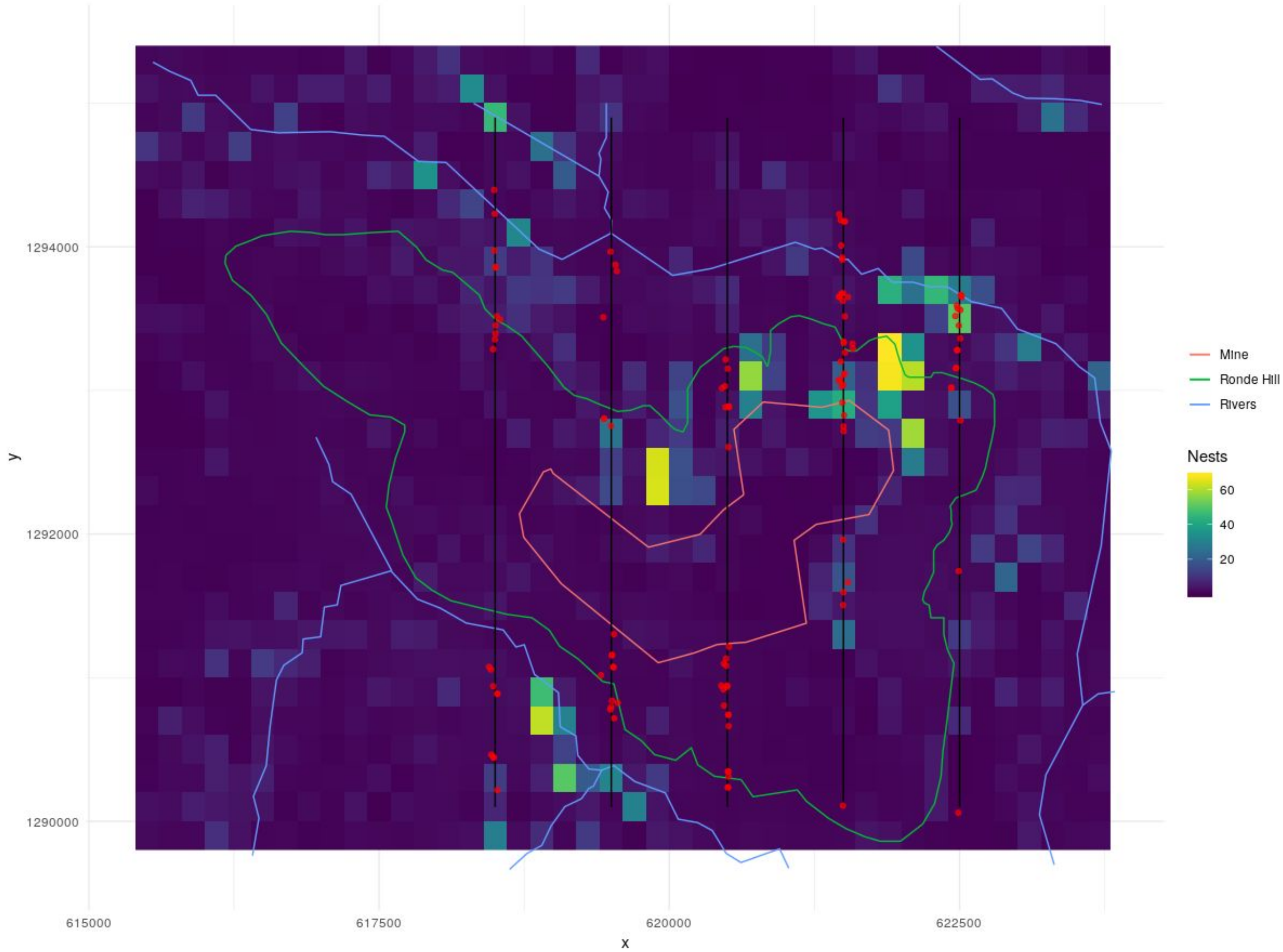



Figure 5 – Predicted abundance of nests overlaid with the location of the transects (black lines), location of clusters of nests (red dots) and future location of the mine (pink line). Ronde hill is shown by the green line and rivers are shown by blue lines.

**Table 1 -** Covariates used in the spatial model (GAM).

| Variables | Description and Units | Source |
|---|---|---|
| Slope | Mean slope (degrees) | ASTER GDEM 2.0 |
| Aspect | Mean aspect (radians) | ASTER GDEM 2.0 |
| Altitude | Mean altitude (m) | ASTER GDEM 2.0 |
| Distance to closest permanent river | Distance (m) | JFCW |
| Distance to closest seasonal river | Distance (m) | JFCW |
| Distance to closest village | Distance to the centroid of the closest village (m) | JFCW |
| Distance to closest road | Distance (m) | JFCW |
| Agriculture | Area (ha) | JFCW |
| Urban | Area (ha) | JFCW |
| Primary Forest | Area (ha) | JFCW |
| Secondary Forest | Area (ha) | JFCW |
| Savanna | Area (ha) | JFCW |
| Land use diversity | Shannon-Wiener diversity Index | - |

**Research Highlights**

- Approximately 18 nest building western chimpanzees inhabit the surroundings of a bauxite deposit in the SW of Guinea-Bissau;

- The construction of a mine can have adverse direct and indirect effects on this population.

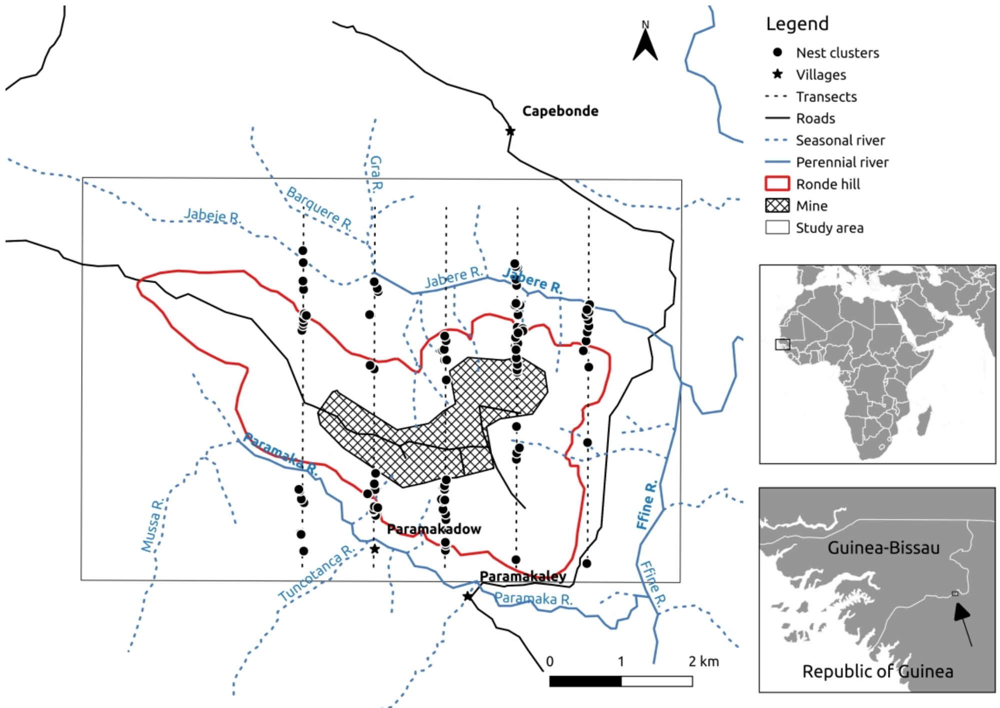



Figure 1

296x209mm (300 x 300 DPI)

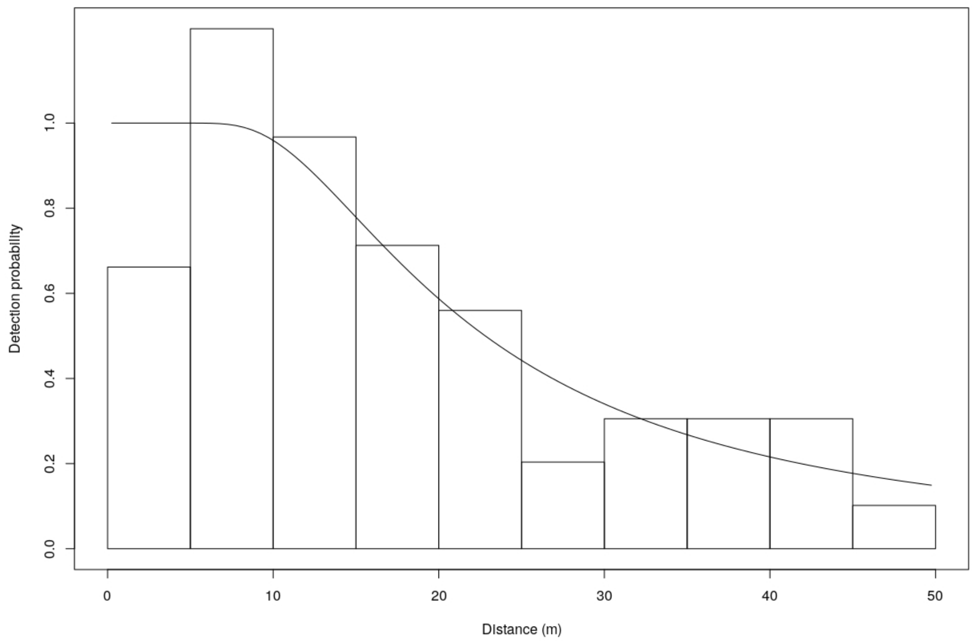



Figure 2

396x285mm (72 x 72 DPI)

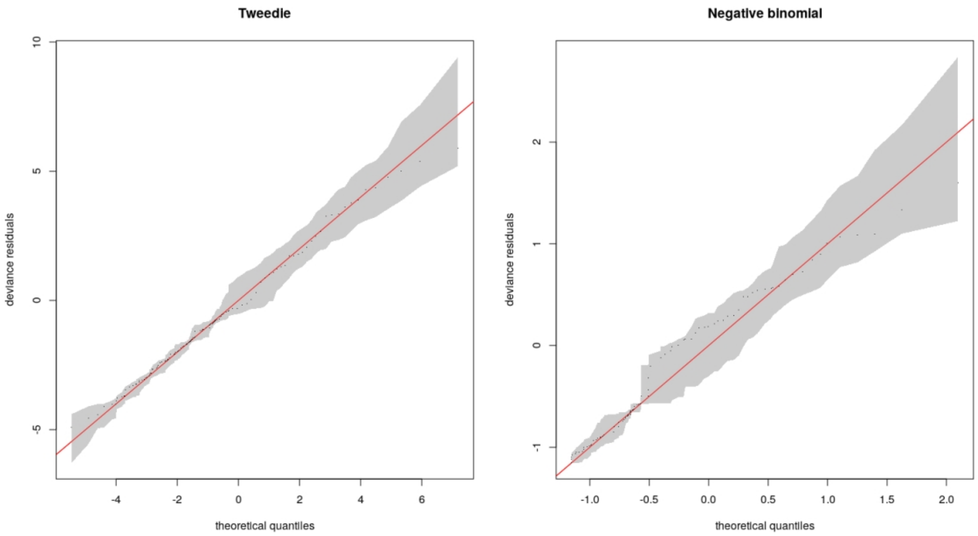



Figure 3

381x213mm (96 x 96 DPI)

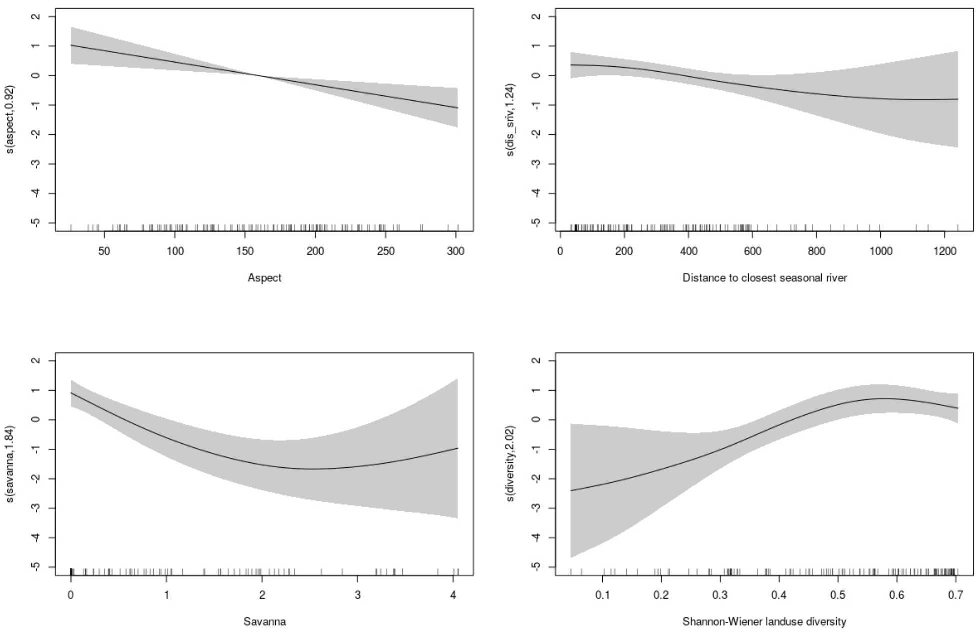



Figure 4

425x293mm (72 x 72 DPI)

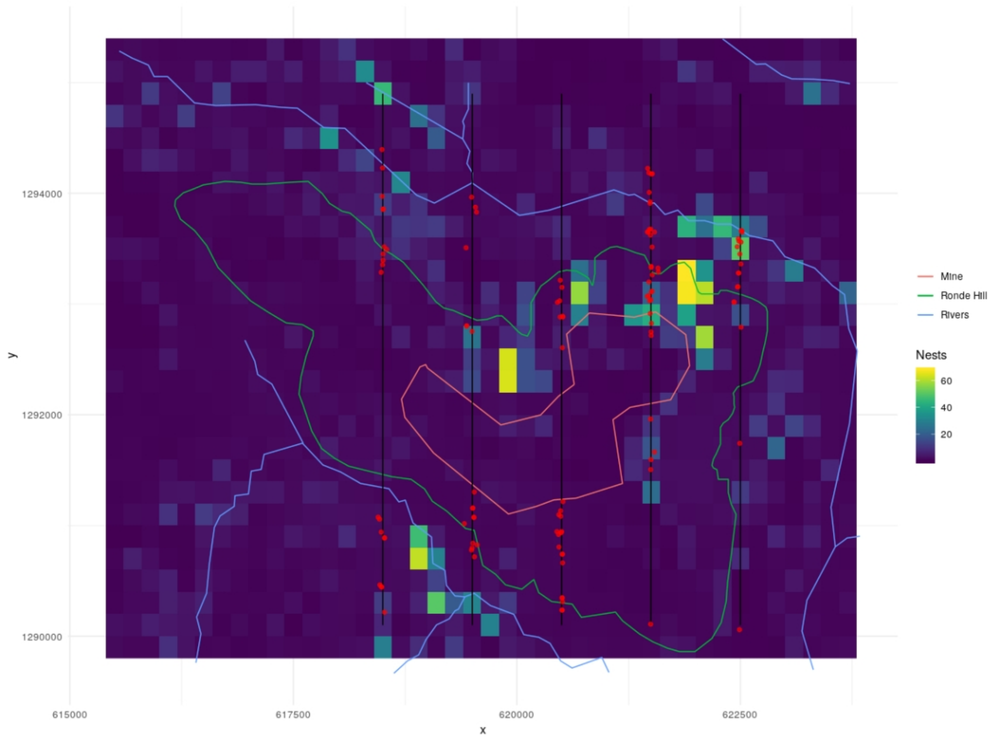



Figure 5

423x313mm (72 x 72 DPI)

# Supplementary information from ‘Density and distribution of western chimpanzees around a bauxite deposit in the Boé Sector, Guinea-Bissau’

*José F. C. Wenceslau*
*Filipe S. Dias*
*Tiago A. Marques*
*David L. Miller*

## 1. Introduction

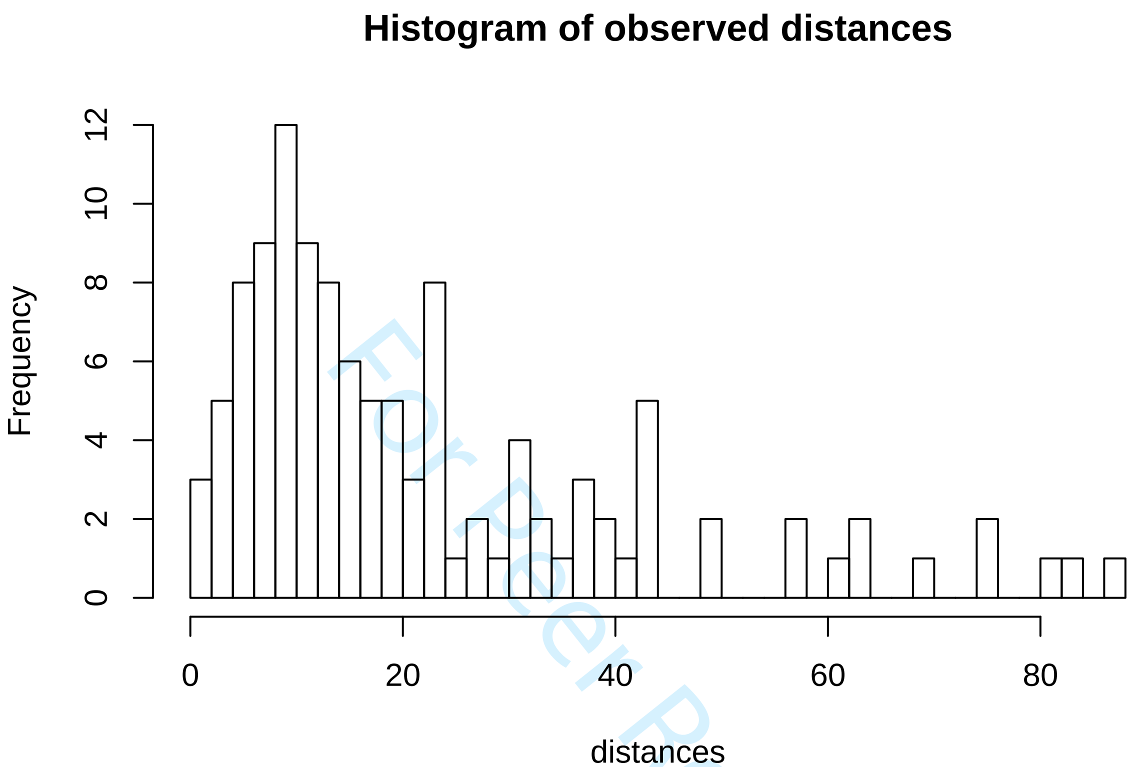
In this document we present the R code we used to generate the results we present and discuss in “Density and distribution of western chimpanzees around a bauxite deposit in the Boé Sector, Guinea-Bissau.”

## 2. Load required packages

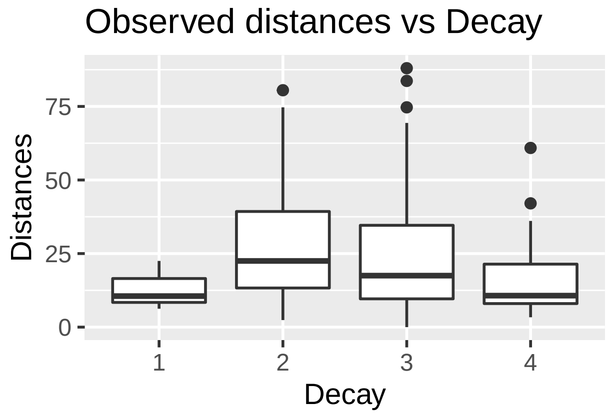
```
library(ggplot2)
library(gridExtra)
library(knitr)
library(mrds)
library(Distance)
library(dsm)
library(tweedie)
library(vegan)
library(viridis)
library(usdm)
```

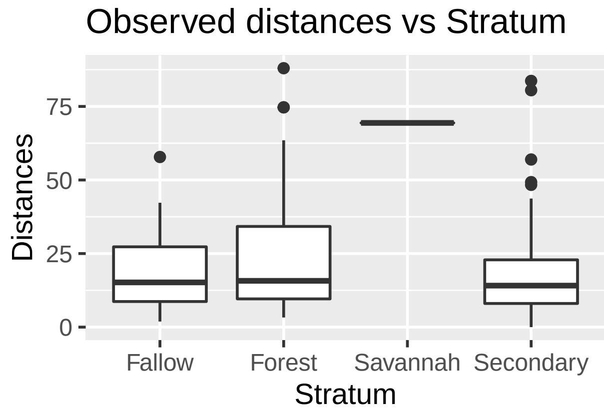
## 2. Exploratory data analysis

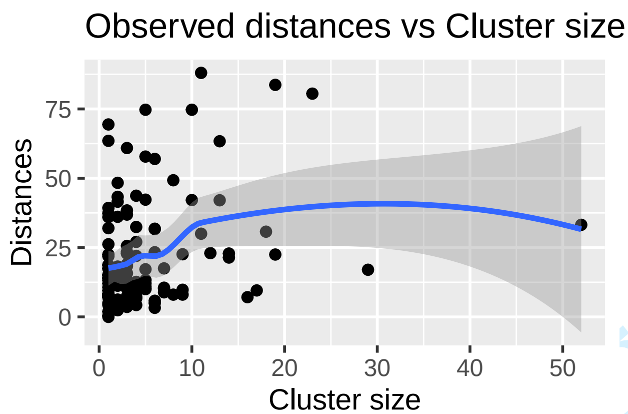
### 2.1 Histogram of observed distances



### 2.2 Do observed distances change as function of covariates?




## 4. Fit detection functions

### 4.1 Conventional distance sampling (CDS)

```
df1c<-ds(data_scnc_clus, truncation=50, key = "unif", adjustment="cos",order=c(2))
df2c<-ds(data_scnc_clus, truncation=50, key = "hn", adjustment="cos",order=c(2))
df3c<-ds(data_scnc_clus, truncation=50, key = "hn", adjustment="herm")
df4c<-ds(data_scnc_clus, truncation=50, key = "hr", adjustment="poly")
```

### 4.2 Multiple-covariate distance sampling (MCDS)

```
#Hazard-rate function
df5c<-ds(data_scnc_clus, truncation=50, key="hr",formula=~size)
df6c<-ds(data_scnc_clus, truncation=50, key="hr",formula=~stratum)
df7c<-ds(data_scnc_clus, truncation=50, key="hr",formula=~decay)
df8c<-ds(data_scnc_clus, truncation=50, key="hr",formula=~size+stratum)
df9c<-ds(data_scnc_clus, truncation=50, key="hr",formula=~decay+stratum)
df10c<-ds(data_scnc_clus, truncation=50, key="hr",formula=~size+decay)
df11c<-ds(data_scnc_clus, truncation=50, key="hr",formula=~size+decay+stratum)

#Half-normal function
df12c<-ds(data_scnc_clus, truncation=50, key="hn",formula=~size)
df13c<-ds(data_scnc_clus, truncation=50, key="hn",formula=~stratum)
```

```
df14c<-ds(data_scnc_clus, truncation=50, key="hn",formula=~decay)
df15c<-ds(data_scnc_clus, truncation=50, key="hn",formula=~size+stratum)
df16c<-ds(data_scnc_clus, truncation=50, key="hn",formula=~decay+stratum)
df17c<-ds(data_scnc_clus, truncation=50, key="hn",formula=~size+decay)
df18c<-ds(data_scnc_clus, truncation=50, key="hn",formula=~size+decay+stratum)
```

### 4.3 Compare candidate detection functions based on AIC and goodness of fit test (Cramer von Mises)

```
df_table<-summarize_ds_models(df1c,df2c,df3c,df4c,df5c,df6c,df7c,df8c,df9c,
                              df10c,df11c,df13c,df14c,df15c,df16c,df18c,sort="AIC")
row.names(df_table)<-c()
kable(df_table[,c("Key function","Formula","C-vM p-value","$\\Delta$AIC")],digits=3,
      caption="Table S1 - Candidate detection functions")
```

Table 1: Table S1 - Candidate detection functions

| Key function | Formula | C-vM p-value | $\Delta$AIC |
|---|---|---|---|
| Hazard-rate | ~size | 0.391 | 0.000 |
| Hazard-rate | ~size + decay | 0.414 | 0.360 |
| Hazard-rate | ~decay | 0.440 | 1.245 |
| Hazard-rate | ~1 | 0.445 | 1.319 |
| Half-normal | ~decay | 0.325 | 1.415 |
| Half-normal | ~1 | 0.346 | 2.344 |
| Uniform with cosine adjustment terms of order 1,2 | NA | 0.353 | 3.020 |
| Half-normal with cosine adjustment term of order 2 | ~1 | 0.305 | 3.477 |
| Hazard-rate | ~size + stratum | 0.404 | 3.764 |
| Hazard-rate | ~size + decay + stratum | 0.397 | 4.087 |
| Half-normal | ~size + decay + stratum | 0.306 | 4.448 |
| Hazard-rate | ~stratum | 0.465 | 5.093 |
| Hazard-rate | ~decay + stratum | 0.422 | 5.101 |
| Half-normal | ~decay + stratum | 0.311 | 5.357 |
| Half-normal | ~size + stratum | 0.316 | 5.510 |
| Half-normal | ~stratum | 0.368 | 6.241 |

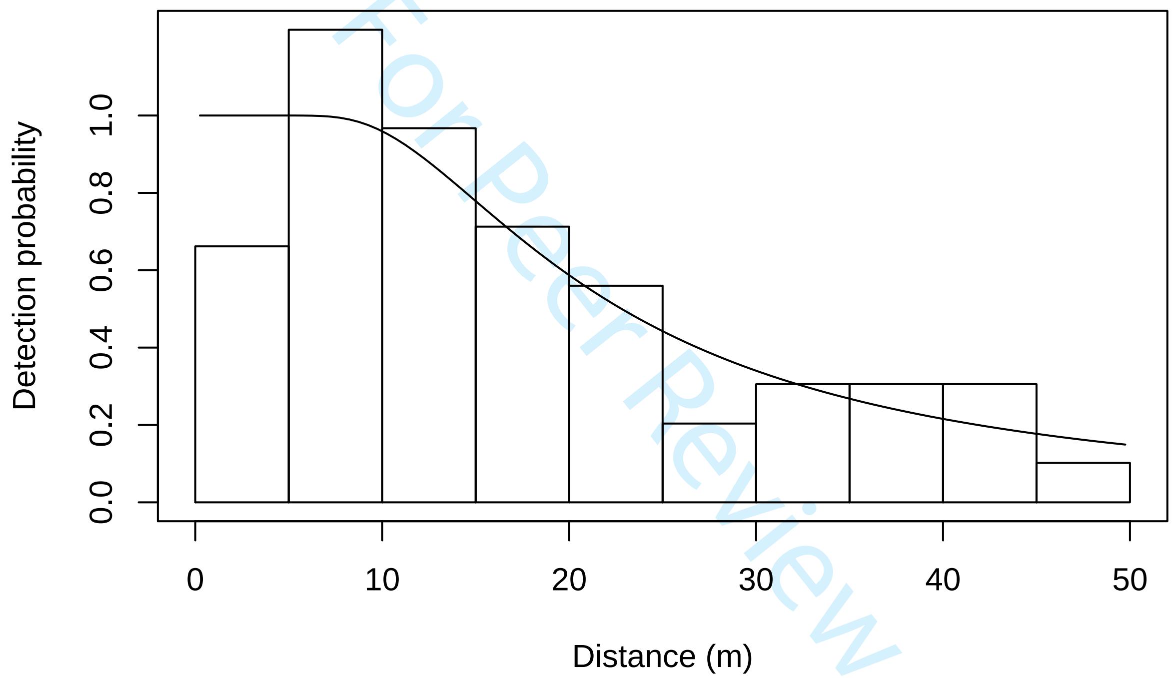
### 4.4 Summary and plot of the selected detection function

```
summary(df5c)
```

```
##
## Summary for distance analysis
## Number of observations :  105
## Distance range         :  0  -  50
##
## Model : Hazard-rate key function
## AIC   : 788.5798
##
## Detection function parameters
## Scale coefficient(s):
##               estimate        se
## (Intercept) 2.74574721 0.29804898
```

```
## size        0.05122658 0.04597067
##
## Shape coefficient(s):
##              estimate        se
## (Intercept) 0.6938415 0.2843393
##
##                          Estimate          SE        CV
## Average p                0.534482  0.06794455 0.1271222
## N in covered region 196.451903 28.28535362 0.1439811
```

```
plot(df5c,breaks=seq(0,50,by=5),showpoints=F, xlab='Distance (m)', cex=1.5)
```






# 5. Density surface models

## 5.1 Covariates

1. `altitude` - mean altitude (m)
2. `slope` - mean slope (%)
3. `zone_type` - conservation (cz) or non-conservation zone (ncz)
4. `aspect` - mean aspect (radians)
5. `dis_priv` - distance to closest permanent river (m)
6. `dis_sriv` - distance to closest seasonal river (m)
7. `dis_road` - distance to closest road (m)
8. `dis_city` - distance to closest city (m)
9. `agriculture` - area of agriculture (ha)
10. `urban` - area of urban areas (ha)

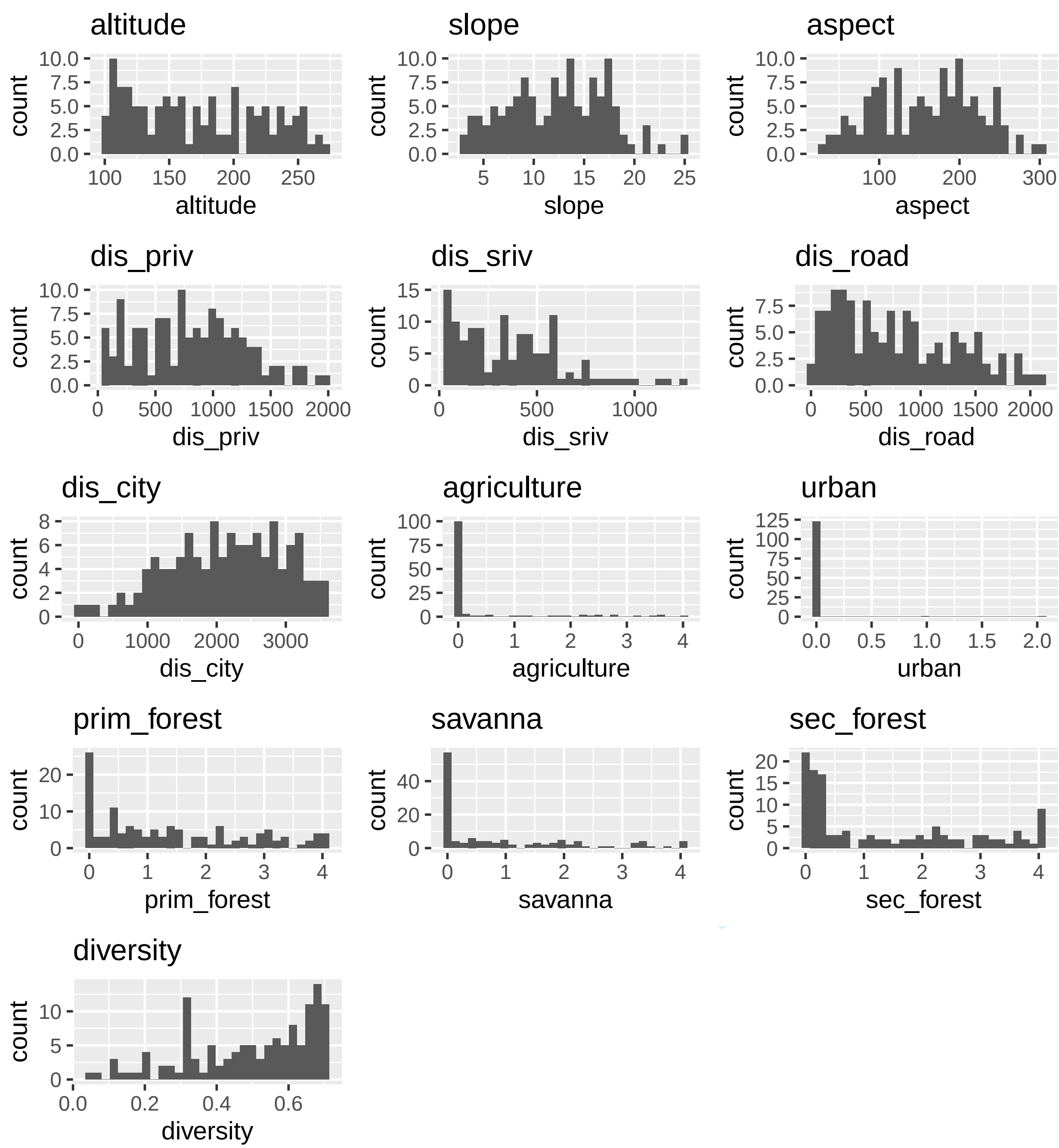
11. `prim_forest` - area of primary forest (ha)
12. `savanna` - area of savanna (ha)
13. `sec_forest` - area of secondary forest (ha)
14. `diversity` - Shannon-Wiener diversity of landuses

### 5.2 Calculate Shannon-Wiener landuse diversity

```
library(vegan)
segment_data$diversity<-diversity(segment_data[,11:15])
```

## 5.3 Explore covariates

### 5.3.1 Histograms with the covariates




### 5.3.2 Assess correlations between variables

```
vifstep(subset(segment_data,select=c(6:17,22)), th=3)
```

```
## 2 variables from the 13 input variables have collinearity problem:
##
## sec_forest dis_road
```

```
##
## After excluding the collinear variables, the linear correlation coefficients ranges between:
## min correlation ( urban ~ aspect ):  -0.006550346
## max correlation ( dis_priv ~ altitude ):  0.6224637
##
## ---------- VIFs of the remained variables --------
##      Variables      VIF
## 1     altitude 2.679982
## 2        slope 1.744131
## 3       aspect 1.295983
## 4     dis_priv 2.884064
## 5     dis_sriv 1.666188
## 6     dis_city 1.930337
## 7  agriculture 1.674478
## 8        urban 1.438926
## 9  prim_forest 1.440064
## 10     savanna 2.263407
## 11   diversity 2.368425
```

## 5.4 Tweedie model

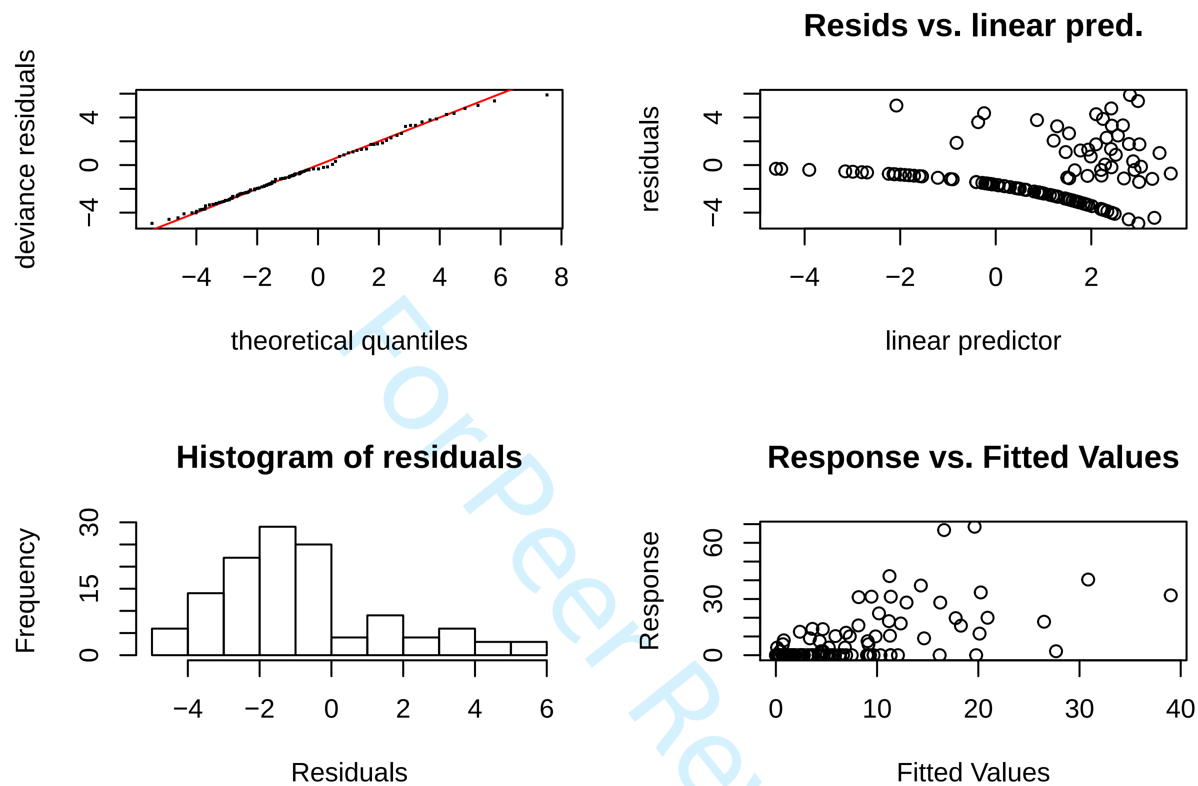
### 5.4.1 Fit the final model

```
model_tw_c<- dsm(Nhat ~ s(aspect)+s(dis_sriv)+s(savanna)+s(diversity),
                 df5c, observation.data=data_scnc_clus,
                 segment.data=segment_data,engine="gam",family=tw(),
                 select=TRUE,method="REML")
summary(model_tw_c)
```

```
##
## Family: Tweedie(p=1.273)
## Link function: log
##
## Formula:
## Nhat ~ s(aspect) + s(dis_sriv) + s(savanna) + s(diversity) +
##     offset(off.set)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -9.0430     0.2508  -36.06   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                 edf Ref.df     F  p-value
## s(aspect)    0.9159      9 1.190 0.000761 ***
## s(dis_sriv)  1.2431      9 0.477 0.034616 *
## s(savanna)   1.8387      9 1.953 5.24e-05 ***
## s(diversity) 2.0221      9 1.259 0.002151 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.355   Deviance explained = 46.9%
## -REML = 226.71  Scale est. = 8.8207    n = 125
```

### 5.4.2 Model validation

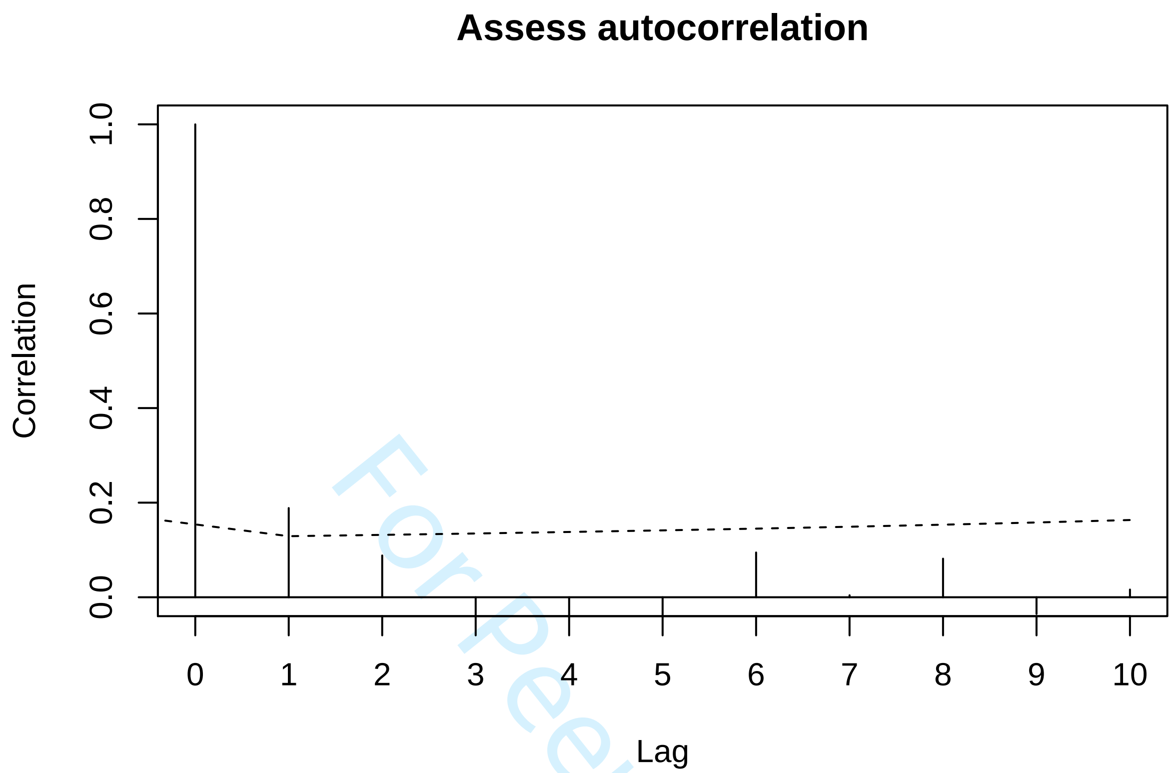
```r
par(mfrow=c(2,2))
gam.check(model_tw_c)
```






```
##
## Method: REML   Optimizer: outer newton
## full convergence after 13 iterations.
## Gradient range [-0.000533106,0.0006441383]
## (score 226.7065 & scale 8.820699).
## Hessian positive definite, eigenvalue range [1.435505e-05,82.05384].
## Model rank =  37 / 37
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                 k'   edf k-index p-value
## s(aspect)    9.000 0.916    0.94    0.81
## s(dis_sriv)  9.000 1.243    0.79    0.14
## s(savanna)   9.000 1.839    0.80    0.14
## s(diversity) 9.000 2.022    0.91    0.67
```

```r
par(mfrow=c(1,1))
dsm.cor(model_tw_c, max.lag = 10,main="Assess autocorrelation")
```

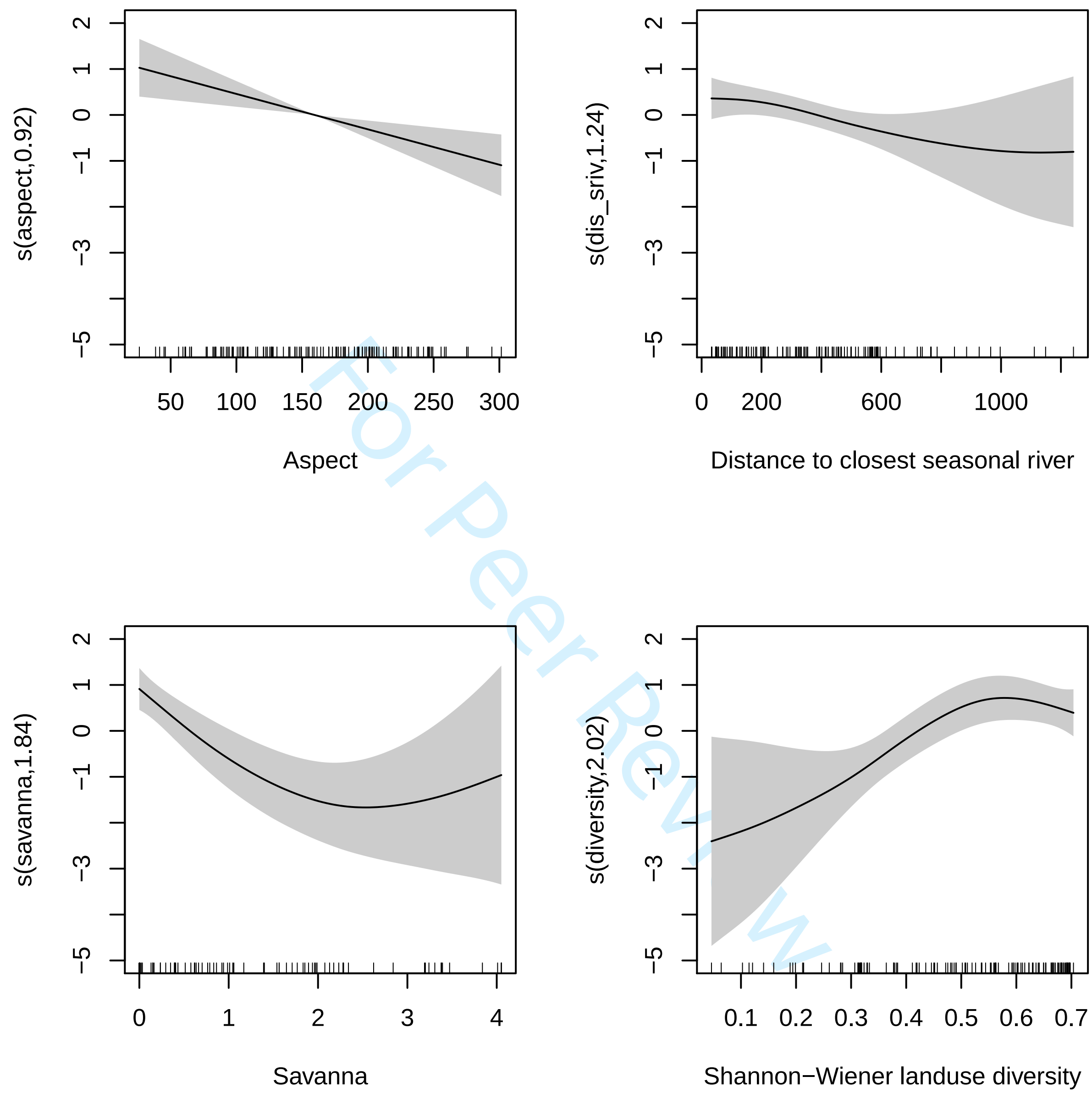



```
concurvity(model_tw_c)
```

```
##                  para s(aspect) s(dis_sriv) s(savanna) s(diversity)
## worst    1.777762e-24 0.6080265   0.5696487  0.6269396    0.6599517
## observed 1.777762e-24 0.2487394   0.4249006  0.4398928    0.5684356
## estimate 1.777762e-24 0.2387085   0.4267187  0.4632259    0.5252859
```

### 5.4.3 Plot smoothers

```
par(mfrow=c(2,2))
plot(model_tw_c, shade=TRUE, ylim=c(-5,2),fig.height=7,select=1,xlab="Aspect")
plot(model_tw_c, shade=TRUE, ylim=c(-5,2),fig.height=7,select=2,xlab="Distance to closest seasonal rive
plot(model_tw_c, shade=TRUE, ylim=c(-5,2),fig.height=7,select=3,xlab="Savanna")
plot(model_tw_c, shade=TRUE, ylim=c(-5,2),fig.height=7,select=4,xlab="Shannon-Wiener landuse diversity")
```






```
par(mfrow=c(1,1))
```

### 5.5 Negative binomial model

#### 5.5.1 Fit the final model

```
model_nb_c<- dsm(Nhat ~ s(slope)+s(aspect)+s(savanna),
                 df5c, observation.data=data_scnc_clus,
                 segment.data=segment_data,engine="gam",family=nb(),
                 select=TRUE,method="REML")
```

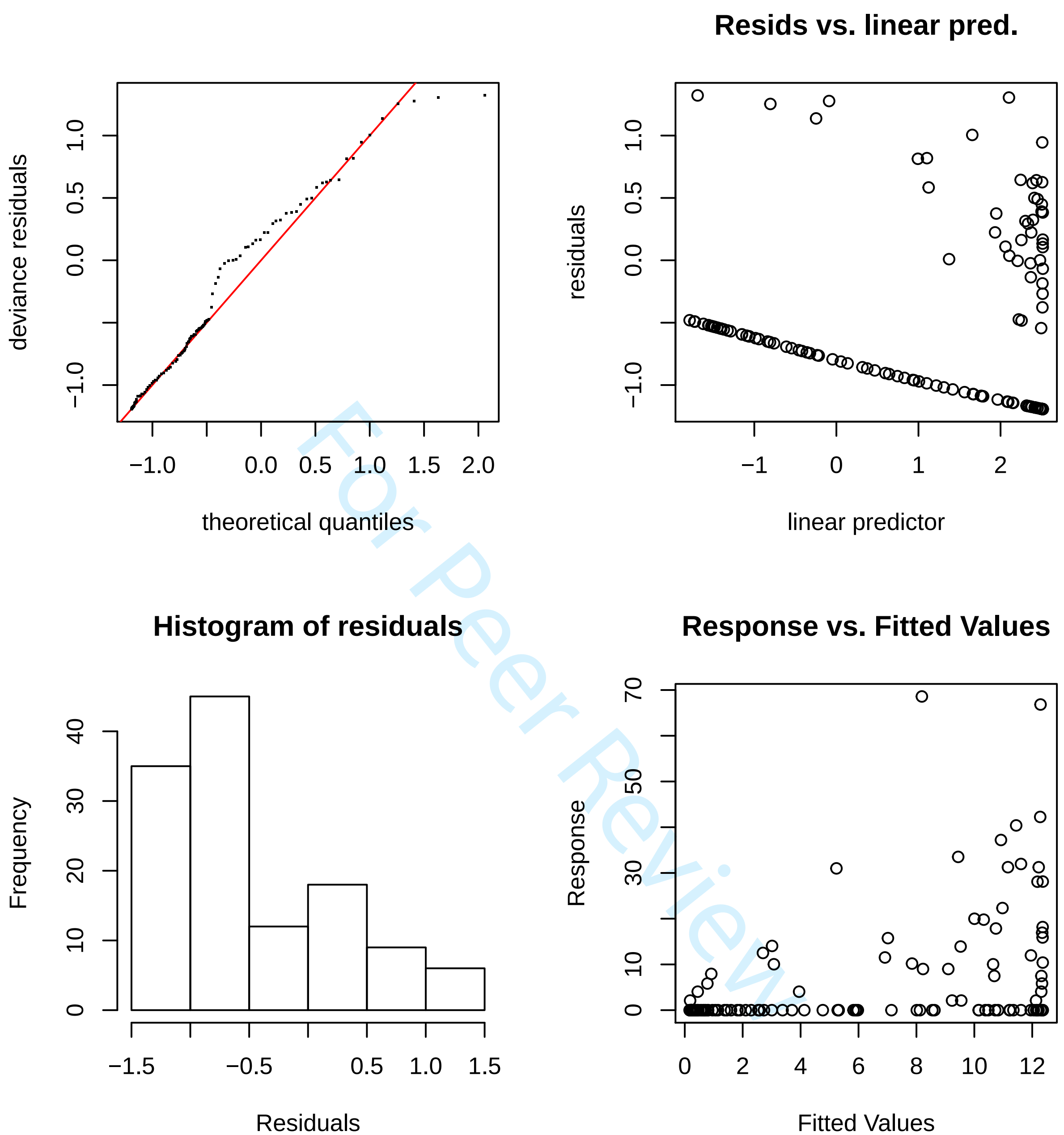
```
## Warning in make.data(response, ddf.obj, segment.data, observation.data, :
## Some observations are outside of detection function truncation!
```

```
summary(model_nb_c)
```

```
##
## Family: Negative Binomial(0.164)
## Link function: log
##
## Formula:
## Nhat ~ s(slope) + s(aspect) + s(savanna) + offset(off.set)
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -8.8602     0.2391  -37.06   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                  edf Ref.df Chi.sq  p-value
## s(slope)   1.1315956      9  2.232    0.128
## s(aspect)  0.0002205      9  0.000    0.597
## s(savanna) 1.9722462      9 24.181 3.99e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.158   Deviance explained = 26.6%
## -REML = 252.32  Scale est. = 1         n = 125
```

### 5.5.2 Model validation

```
par(mfrow=c(2,2))
gam.check(model_nb_c)
```

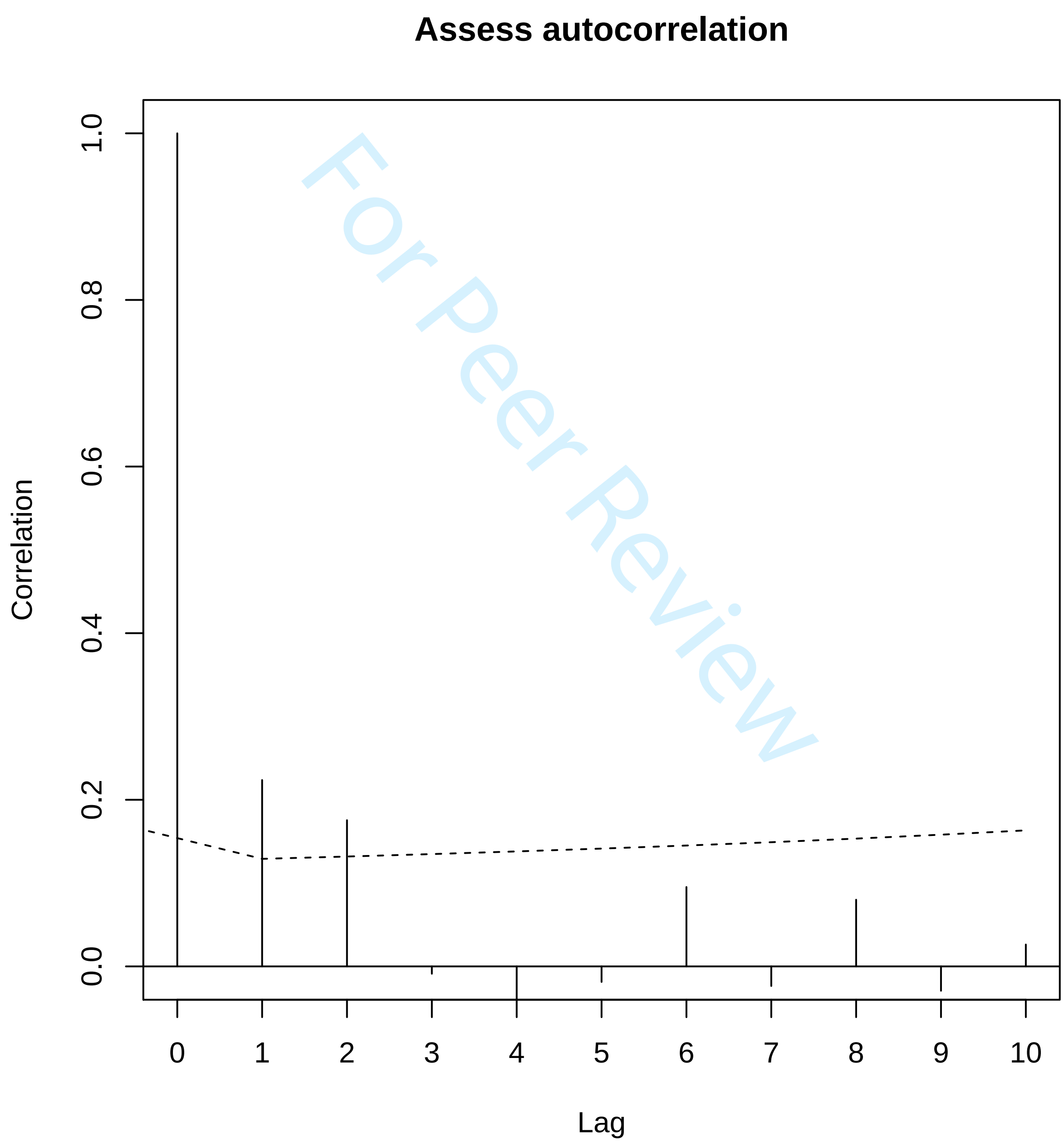





```
##
## Method: REML   Optimizer: outer newton
## full convergence after 11 iterations.
## Gradient range [-8.869582e-05,9.496902e-05]
## (score 252.3186 & scale 1).
## Hessian positive definite, eigenvalue range [6.644452e-06,24.35469].
## Model rank =  28 / 28
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
```

```
##                  k'     edf  k-index p-value
## s(slope)    9.00000 1.13160     0.69   0.435
## s(aspect)   9.00000 0.00022     0.74   0.770
## s(savanna)  9.00000 1.97225     0.54   0.005 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(1,1))
dsm.cor(model_nb_c, max.lag = 10,main="Assess autocorrelation")
```

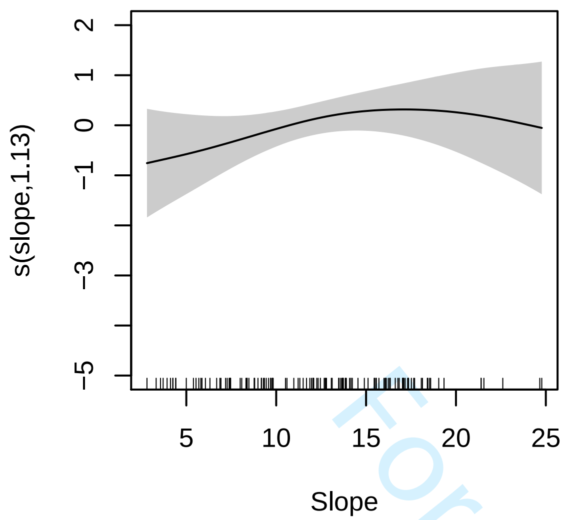


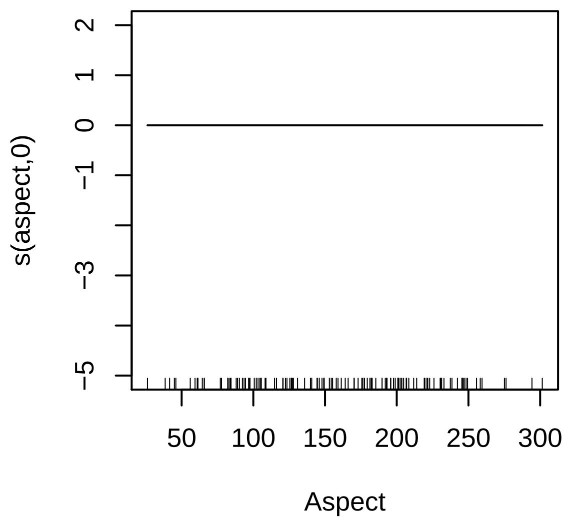
```
concurvity(model_nb_c)
```

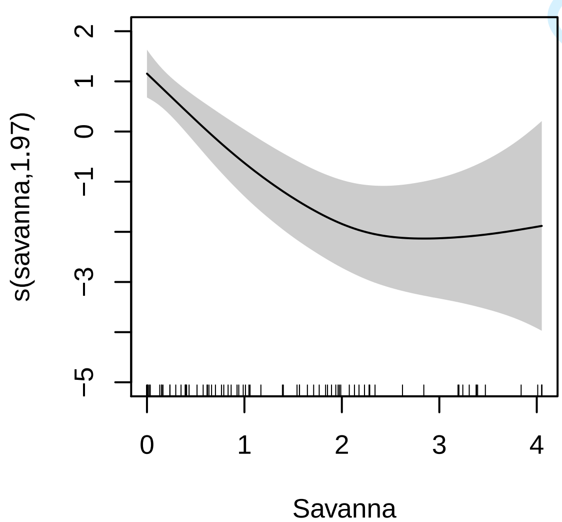
```
##                  para  s(slope) s(aspect) s(savanna)
## worst    9.027777e-25 0.5553578 0.3767250  0.5460532
## observed 9.027777e-25 0.4459200 0.1783605  0.4180740
## estimate 9.027777e-25 0.3620254 0.1609830  0.4809394
```

### 5.5.3 Plot smoothers

```
par(mfrow=c(2,2))
plot(model_nb_c, shade=TRUE, ylim=c(-5,2),fig.height=7,select=1,xlab="Slope")
plot(model_nb_c, shade=TRUE, ylim=c(-5,2),fig.height=7,select=2,xlab="Aspect")
plot(model_nb_c, shade=TRUE, ylim=c(-5,2),fig.height=7,select=3,xlab="Savanna")
par(mfrow=c(1,1))
```



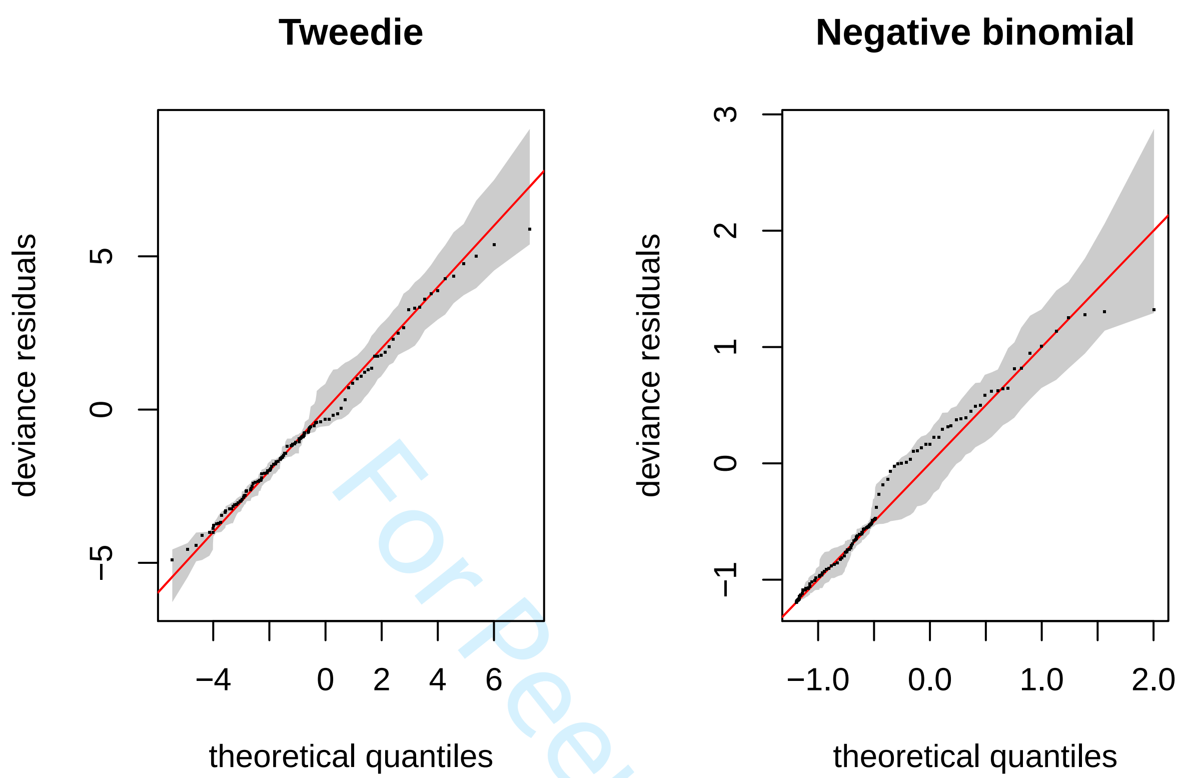






### 5.6 Which model should we select?

```
par(mfrow=c(1,2))
qq.gam(model_tw_c,rep=100,main="Tweedie")
qq.gam(model_nb_c,rep=100,main="Negative binomial")
```




```
par(mfrow=c(1,1))
```

This plot shows a comparison of models with Tweedie (left) and negative binomial (right) response distributions by quantile-quantile plots. Good fit is indicated by agreement between observed and fitted (residual) quantiles (i.e., points being close to the red line). 90% reference bands are shown in grey allowing judgement of the deviation from the line. The negative binomial points fall further away from the red line than those for the Tweedie, indicating model misspecification.

# 6. Model predictions

## 6.1 Calculate offset

```
off.set <- (200 * 200) #grid is 200 m x 200 m
```

## 6.2 Predictions from the Tweedie model

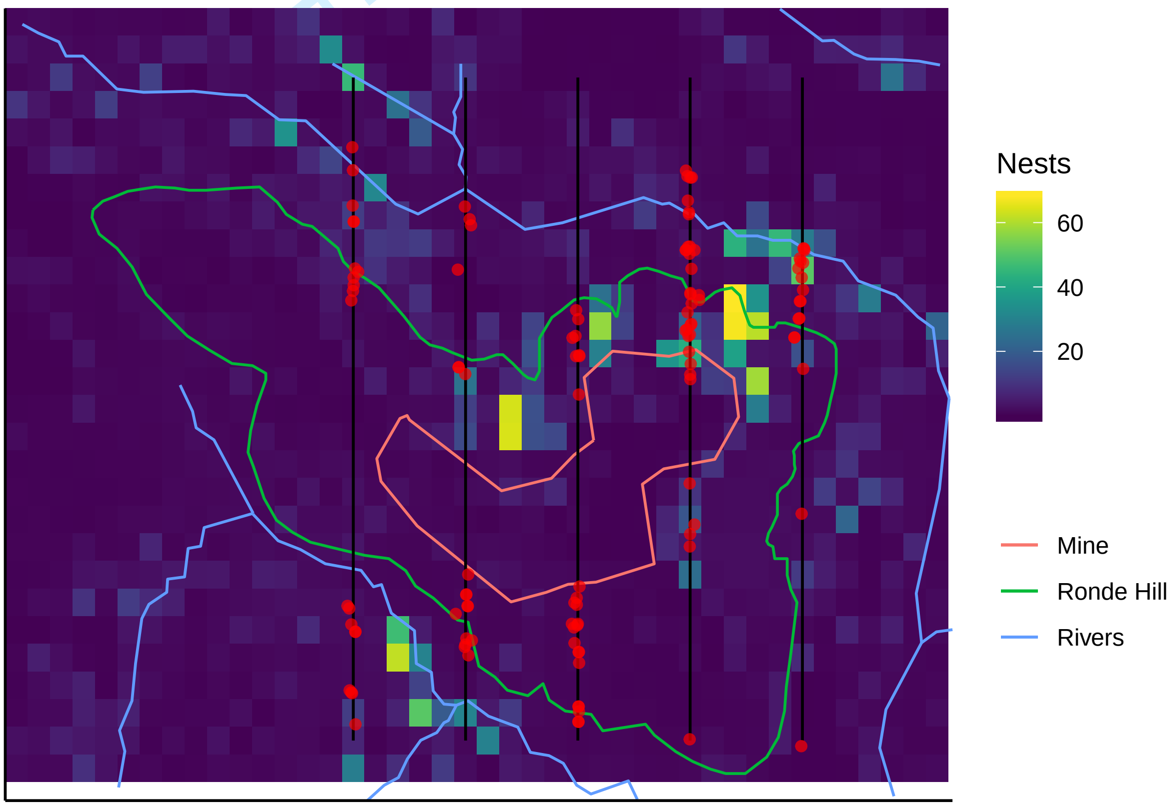
### 6.2.1 Calculate predicted abundances

```
model_tw.pred_c <- predict(model_tw_c, preddata, off.set)
preddata$TW_ab_c<-unname(model_tw.pred_c)
```

### 6.2.2 Plot predicted abundances alongside transects and clusters of nests

```
p<-ggplot(preddata, aes(x, y))+theme_minimal()
p<-p + geom_raster(aes(fill = TW_ab_c))
```

```
p<-p + scale_fill_viridis(name="Nests")
p<-p + geom_path(data=mine, aes(long, lat, colour="black"))
p<-p + geom_path(aes(x=POINT_X, y=POINT_Y, colour = "brown4"), data = Points_areas[26:185,])
p<-p + geom_path(aes(x=POINT_X, y=POINT_Y, group = ORIG_FID, colour = "darkcyan"),
                 data = Points_river_roads[1:121,])
p<-p + geom_line(aes(x, y, group = Transect.Label), data = segment_data)
p<-p + geom_point(aes(x=x, y=y), data = data_scnc_clus, colour = "red",
                  alpha = I(0.7))
p<-p+scale_color_hue(labels = c("Mine", "Ronde Hill","Rivers","Transect"))+
  labs(color='')
p<-p+theme(panel.grid.major = element_blank(),panel.grid.minor = element_blank(),
           axis.line = element_line(colour = "black"))
p<-p+theme(axis.text.x=element_blank(),axis.text.y=element_blank())+labs(x="",y="")+
  coord_cartesian(expand=FALSE)
p
```




### 6.2.3 Calculate prediction variances

```
model_tw_var_c<- dsm.var.gam(model_tw_c, pred.data = preddata, off.set = off.set)
summary(model_tw_var_c)
```

```
## Summary of uncertainty in a density surface model calculated
##  analytically for GAM, with delta method
##
## Approximate asymptotic confidence interval:
##     2.5%     Mean    97.5%
```

```
## 2268.237 3877.604 6628.855
## (Using log-Normal approximation)
##
## Point estimate                 : 3877.604
## CV of detection function       : 0.1271222
## CV from GAM                    : 0.2481
## Total standard error           : 1081.014
## Total coefficient of variation : 0.2788
```

# 7. Calculate density and abundance of nest building chimpanzees with the Tweedie model

To calculate the density of chimpanzees we use the following formula:

D_weaned_chimpanzee=D_nests/(r*t)

where "r" is the estimated rate of nest production per individual per day estimated to be 1.09 nests/individual/day by Plumptre & Reynolds (1997) and "t" is the mean life of a nest estimated to be 194 days by Fleury-Brugiere & Brugiere (2010).



Following this formula, the estimated number of weaned chimpanzees in the study area is:

```
weaned_chimps<- as.numeric(model_tw_var_c$pred)/(1.09*194)
print(weaned_chimps)
```

```
## [1] 18.33729
```

Now, we calculate the 95% confidence intervals for the number of weaned chimpanzees in the study area using the upper and lower bounds of the estimated number of nests (see above):

```
weaned_chimps_upper <- 6628.829/(1.09*194)
print(weaned_chimps_upper)
```

```
## [1] 31.34791
```

```
weaned_chimps_lower <- 2268.236/(1.09*194)
print(weaned_chimps_lower)
```

```
## [1] 10.72655
```

Considering the study area covers 47.04 squared kilometers, the number of chimpanzees per squared kilometer and the corresponding 95% confidence interval is:

```
estimate<- c(weaned_chimps, weaned_chimps_lower, weaned_chimps_upper)
final_value<- estimate/47.04
print(final_value)
```

```
## [1] 0.3898234 0.2280304 0.6664096
```